DOCUMENT RESUME

ED 473 748 CG 032 202

AUTHOR	McCutcheon, Lynn E., Ed.
TITLE	North American Journal of Psychology, 2002.
ISSN	ISSN-1527-7143
PUB DATE	2002-00-00
NOTE	507p.; Published in March, June and December.
AVAILABLE FROM	NAJP, 240 Harbor Dr., Winter Garden, FL 34787 ($30 per annual individual subscription). Tel: 407-877-8364.
PUB TYPE	Collected Works - Serials (022)
JOURNAL CIT	North American Journal of Psychology; v4 n1-3 2002
EDRS PRICE	EDRS Price MF02/PC21 Plus Postage.
DESCRIPTORS	*Psychology; *Research Tools; *Scholarly Journals; *Social Science Research

ABSTRACT

"North American Journal of Psychology" publishes scientific papers of general interest to psychologists and other social scientists. Articles included in volume 4 issue 1 (March 2002) are: "An Interview with Kimmo Lehtonen: Music Therapy with Adolescents"; "The Relationship of Verbal-Nonverbal Incongruence to Communication Mismatches in Married Couples"; "The Effects of Confidence Perception of Test-Taking Skills on Performance"; "Relationships between Dimensions of Adult Attachment and Empathy"; "Programs for Young Fathers: Essential Components and Evaluation Issues"; "Meaning in Life and Positive and Negative Well-Being"; "Validity of the Wonderlic Personnel Test as a Measure of Fluid or Crystallized Intelligence: Implications for Career Assessment"; "Child Sexual Abuse: A Review of Definitions, Instrumentation, and Symptomology." Articles included in volume 4 issue 2 (July 2002) are: "An Interview with Arnold A. Lazarus"; "Depression, Drug Use, and Gender Differences among Students at a Religious University"; "Body-Image Assessment: A Review and Evaluation of a New Computer-Aided Measurement Technique"; "A Short Scale of Family Atmosphere (SOFA): Development and Psychometric Evaluation"; "An Interview with Mark Ylvisaker about Students with Traumatic Brain Energy"; "Delivering Human Services to native Americans with Disabilities: Cultural Variables and Recommendations"; "On Differentiating Major Depression from Chronic Sadness: A Commentary." Articles included in volume 4 issue 3 (December 2002) are: "Girls with ADHD and Associated Behavioral Problems: Patterns of Comorbidity"; "An Interview with Albert Ellis about Rational Emotive Behavior Therapy"; "Counselor-Client Matching on Ethnicity, Gender, and Language: Implications for Counseling School-Aged Children"; "Parent Discipline Scale: Parental Discipline Styles as a Function of Transgressor Type"; "The Effect of Autobiographical Writing in the Subjective Well-Being of Older Adults"; "Psychology Training Regarding HIV/AIDS Revisited"; "Attachment Styles, View of Self, and Negative Affect"; "Young Females' Perceptions of the Impact of a Sexual Abuse Experience: A Focus Group Approach"; "An Interview with Sonia Nieto about Multiculturalism." (GCP)

Reproductions supplied by EDRS are the best that can be made from the original document.

ED 473 748

North American Journal of Psychology
Volume 4
2002

Lynn McCutcheon, Editor

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

McCutcheon

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

BEST COPY AVAILABLE

CG032202

North American Journal of Psychology

Vol. 4, No. 1 **2002**

3

Advisory Editorial Board

ADRIAN FURNHAM
Psychology
University College London

DOUGLAS JACKSON
Psychology
University of Western Ontario

JAMES M. JONES
Psychology
University of Delaware

HEIDI KELLER
Psychology
University of Osnabrueck

MANUEL LONDON
Psychology
SUNY, Stony Brook

KAREN LYNESS
Management
Baruch College, CUNY

ROBERT PASNAK
Psychology
George Mason University

RICHARD RYAN
Psychology
University of Rochester

PHILIP ZIMBARDO
Psychology
Stanford University

Editor
LYNN E. McCUTCHEON

How to Subscribe to The *North American* Journal of Psychology

All subscriptions are on a calendar year basis only. If you order late in the year, the earlier issue will be sent to you. Institutional rate: $65. Individual rate: $30. Add $7 for Canadian subscriptions. Add $8 for subscriptions outside North America. *North American* Journal of Psychology is published in March, June and December.

Now visit our website at http://najp.8m.com

North American Journal of Psychology

March, 2002 ISSN 1527-7143 Vol. 4, No. 1

"In What Font Color is Bill Cosby's Name Written?" Automatic Racial Categorization in a Stroop Task 1
Jerzy J. Karylowski, Michael Motes, Danielle Curry & Diana Van Liempd

An Interview with Kimmo Lehtonen: Music Therapy with Adolescents 13
Kimmo Lehtonen & Michael F. Shaughnessy

The Relationship Verbal-Nonverbal Incongruence to Communication Mismatches in Married Couples 21
Amy Van Buren

The Effects of Confidence and Perception of Test-taking Skills on Performance 37
Lisa F. Smith

Stimulus Load and Age in Face Recognition: A Comparison of Children and Adults 51
Mitchell M. Metzger

Relationships Between Dimensions of Adult Attachment and Empathy 63
Jeffrey A. Joireman, Tami Lynn Needham, & Amy-Lynn Cummings

Programs for Young Fathers: Essential Components and Evaluation Issues 81
Peggy B. Smith, Ruth S. Buzi, & Maxine L. Weinman

Meaning in Life and Positive and Negative Well-Being 93
Evelyn D. Scannell, Felicity C. L. Allen, & Janet Burton

Validity of the Wonderlic Personnel Test as a Measure of Fluid or Crystallized Intelligence: Implications for Career Assessment 113
Nancy L. Bell, T. Darin Matthews, Kerry S. Lassiter, & J. Patrick Leverett

Parental Acceptance-Rejection of the Disabled Children in Northern Pakistan 121
Zafar Afaq Ansari

Sleep Positions and Personality: An Empirical Study 129
Michael Schredl

Factors Influencing Opinions about Urban Growth and Development: Measuring Resistance to Urban Development 133
James T. Kitchens, Larry Powell, & Davis Cookson

Eyewitness Testimony and the Jury Verdict 143
Jennifer N. Sigler & James V. Couch

Child Sexual Abuse: A Review of Definitions, Instrumentation, and Symptomology 149
Karen Mannon & Gloria Leitschuh

Blood Types and Athletic Performance 161
Gordon W. Russell & Masao Ohmura

Predicting Magnitude Estimates of Drink Strength 165
Sharolyn R. Higgs, Stuart J. McKelvie, & Lionel G. Standing

Editor's Comments

Let me begin by apologizing for the omission of an important figure in the fine article by Timothy Moore that appeared in the last issue. Thanks again to all the reviewers, all of whom did a fine job of critiquing papers promptly and with great insight. Without our reviewers, this journal would simply not have the quality it has today.

Most of our readers will readily identify the *Journal of Personality and Social Psychology* as one of the most prestigious of its kind. In the February, 2002 issue of *Monitor on Psychology,* new JPSP editor John Dovidio foresaw changes during his tenure. Specifically, he pointed out the need to make JPSP's turnaround time faster, the wish to help authors instead of turning them away, and the desire to publish articles that are brief. In other words, Dovidio wants his journal to become more like NAJP. We applaud his willingness to make JPSP more "friendly," and we encourage other editors to do the same.

Lynn E. McCutcheon, editor

7

"In What Font Color Is Bill Cosby's Name Written?": Automatic Racial Categorization in a Stroop Task

Jerzy J. Karylowski
University of North Florida and Instytut Psychologii Polskiej Akademii Nauk

Michael A. Motes, Danielle Curry, & Diana Van Liempd
University of North Florida

We used a modification of the Stroop (1935) color-naming task to investigate spontaneous, unintentional categorization based on race. Participants were presented with names of African American and Caucasian celebrities. The names were written either in black, white, green, or blue font against a background of a different color. The experimental task was to name the font color. We reasoned that because *black* and *white* are used in colloquial English as both color-labels and race-labels, spontaneous categorization of African American and Caucasian experimental targets by race should result in Stroop-like effects on a color naming task. Specifically, we expected that identifying font color as black would be faster when a name of an African-American target is presented in that font than when a name of a Caucasian target is presented. Similarly, identifying white font color should be faster when a name of a Caucasian target is presented in that font. Our results supported these predictions.

"The problem of the Twentieth Century is the problem of the color-line." declared W. E. B. Du Bois (1903) at the century's beginning. "A line runs through the heart of America, cuts across . . . . the high, curved dais of the Etowah County Commission in Alabama when one black member sits with five whites, ... encircles the 'black tables' when African-Americans cluster together during meals at Princeton University, ... intertwines itself through police departments . . . and jury rooms, . . . through television and radio" writes David A. Shipler (1997, pp. 3-4) at the century's end. The two quotes, separated by the 1954 Supreme Court anti-segregation ruling in the *Brown v. Board of Education*, provide an

---

Author info: Correspondence should be sent to: Jerzy J. Karylowski, Department of Psychology, University of North Florida, 4567 St. Johns Bluff Road, South, Jacksonville, Fl 32224-2645; e-mail: jurek@unf.edu.

North American Journal of Psychology, 2002, Vol. 4, No. 1, 1-12.

© NAJP

eloquent illustration of the powerful and pervasive nature of race as a social category.

Not surprisingly, race is often considered in social psychological theory and research as one of the principal categories that perceivers routinely apply in processing information about people (Brewer, 1988; Fiske & Neuberg, 1990; Hamilton & Sherman, 1994; Kunda and Thagard, 1996; Smith & Zarate 1992). Furthermore, it is assumed that categorization based on race may underlie the activation of racial stereotypes (Bargh & Pietromonaco, 1982; Devine, 1989; Lepore & Brown, 1997; Spencer, Fein, Wolfe, Fong, & Dunn, 1998; Wittenbrink, Judd, & Park, 1997). In fact, it has been argued that activation of such stereotypes occurs automatically and inevitably upon a mere exposure to a category label or to other stimuli associated with the category (Bargh & Pietromonaco, 1982; Devine, 1989; Wittenbrink, Judd, & Park, 1997 but see, Gilbert & Hixon, 1991; Spencer, et al., 1998).

In contrast to extensive research on the contribution of automatic and controlled processes to the activation and application of racial stereotypes (see Bodenhausen & Macrae, 1998; Hamilton & Sherman, 1994 for reviews), the question of automaticity of its presumed precursor, racial categorization, received considerably less attention. For instance, in a recently published, multi-staged model of stereotyping, Bodenhausen and Macrae (1998) acknowledge that "categorization and stereotyping can certainly be dissociated empirically and theoretically." Nevertheless, they include categorization and automatic stereotype activation in a single, initial, stage of their model and provide very little elaboration regarding the automatic nature of the categorization process and its role as a precursor of stereotype activation. Such relative neglect might be at least partially due to the methodological challenges of studying race-based categorization as an automatic process separate from activation of racial stereotypes. Most of the relevant work in this area (e.g., Biernat & Vescio, 1993; Hewstone, Hantzi, & Johnston, 1991; Stangor, Lynch, Duan, & Glass, 1992) utilized a variation of the *Who Said What?* paradigm (Taylor, Fiske, Etcoff, & Ruderman, 1978). In a typical experiment, after listening to a group discussion, participants were asked to recall which statement was made by which group member. Within-category errors occurred if a statement was assigned to a wrong person belonging to the same social category as the speaker of the statement (eg, a statement made by one African-American was assigned to another African-American). Even though such within-category errors may be indicative of social categorization, this paradigm is not well equipped to distinguish between, presumably spontaneous, unintentional, categorization occurring during encoding and strategic, performancemaximizing categorization occurring during the retrieval (cf. Klauer & Wegener, 1998). In other studies participants were explicitly asked to perform a categorization task (Stroessner, 1996; Zarate & Smith, 1990) or to make similarity judgments (e.g., Fazio & Dunton, 1997), neither task particularly well suited to assess categorization as an unintentional process.

Perhaps the strongest evidence to date for the spontaneous, unintentional nature of racial categorization comes from experiments by Spencer, et al., (1998, Experiments 2 & 3).[1] Those experiments demonstrated that (under additional conditions) subliminal exposure to race category exemplars (photographs of African American targets) results in activation of concepts associated with a stereotypical view of African Americans (e.g., *dangerous* , *janitor*, *welfare*). Such spontaneous activation of stereotype-related concepts strongly suggests that race-category labels have also been activated. However, because their research did not measure categorization, the evidence is only indirect.[2]

The present experiment utilized a modification of the Stroop (1935) color-naming task to study automatic categorization by race. As in the original Stroop task, participants were presented with verbal stimuli written in fonts of different colors and the experimental task was to name the color of the font. However, instead of verbal color labels (such as *red*, *green*, or *blue*) used in the original Stroop task, the present task employed names of familiar persons (popular actors, singers, and TV personalities), some of them African Americans and some Caucasians. The names were written either in black, white, green, or blue font against a background of a different color. We were particularly interested in comparisons involving two font colors: black and white (the remaining two colors were chosen arbitrarily). We reasoned that because *black* and *white* are used in English as both color-labels and race-labels, spontaneous categorization of African-American and Caucasian experimental targets by race should result in Stroop-like effects on a color naming task. Specifically, we expected that identifying black font color should be faster when a name of an African-American target is presented in that color than when a name of a Caucasian target is presented. Similarly, identifying white font color should be faster when a name of a Caucasian target is presented in that color.

As is usually the case in the color-naming task experiments (see MacLeod, 1991 for a review), participants were instructed to ignore the content of the verbal stimuli (in the present experiment names of African American and Caucasian celebrities) and to concentrate on the font color only. Thus, there was no task-related advantage to be gained from categorizing targets by race or from any other kind of semantic processing of the names presented. We reasoned that finding the predicted Stroop-like effects for names of familiar African American and Caucasian targets would provide indication of the spontaneous, unintentional nature of racial categorization.

The color-naming task was followed by an unexpected free recall task in which participants were asked to list all target names that they could remember. The main purpose of this auxiliary task was to generate data that could be examined for evidence of clustering not only by race but also by sex of the target, thus, potentially, enabling comparisons between relative strengths of the two principles of social categorization.

We realized that, much like within-category confusions in the *Who Said What?* paradigm (Taylor, Fiske, Etcoff, & Ruderman, 1978), clustering in free recall is conceptually ambiguous as a measure of social categorization. Specifically, clustering may reflect not only spontaneous categorization processes occurring during encoding but also, theoretically less interesting, strategic efforts to maximize performance during the retrieval. Furthermore, given the fact both the experimental instructions and the demands of the color naming task discouraged intentional semantic encoding (reading) of target names, we could expect relatively low performance on the free recall task, thus making clustering measures less reliable. In choosing to include the free recall task in spite of those problems, we were swayed by the fact that such task could be easily incorporated into our design. Furthermore, because it was performed after the completion of the color naming task, no interference with the color naming task could occur.

METHOD

Participants

Ninety-four Caucasian undergraduate students (24 males and 70 females) volunteered to participate in the experiment in exchange for extra credit. Data from 32 non-Caucasian students who also volunteered were not included in the final analysis.[3]

Procedure

All instructions and stimuli were presented on a CRT screen. The experiment consisted of 308 trials; the first 20 of which were practice trials. A *Get Ready* prompt appeared for 10s or until the participant pressed a key before each trial. For each trial the program presented a target name in one of four colors (white, black, green, blue) against a background of a 290 X 200 pixels, differently-colored patch (again: white, black, green, or blue). The patch was always centered on a gray screen and the name itself appeared in one of three horizontal locations within the patch (row 218, 256, or 296 on a 640 x 400 pixel display). The experimental task was to identify the color of the font in which the name was written and enter a response via labeled response-keys on the response box, (the middle key of the 5-key box was not used). A maximum of 2 seconds was allowed for each response. The assignment of the two crucial response keys: the *black* key and the *white* key was counterbalanced across participants.

Names of 24 well known, male and female popular-culture celebrities, 12 African-Americans (Louis Armstrong, Bill Cosby, Whoopi Goldberg, Arsenio Hall, Whitney Houston, Janet Jackson, Eddie Murphy, Diana Ross, Tina Turner, Denzel Washington, Montel Williams, Oprah Winfrey) and 12 Caucasians (Roseanne Barr, Garth Brooks, Cindy Crawford, Harrison Ford, Mel Gibson, Jay Leno, Demi Moore, Rosie O'Donnell, Brad Pitt, Jerry Seinfield, Barbara Streisand, Elizabeth Taylor) were used as experimental targets[4]. We paired each of the 24 names with each of the four font colors and each of the 96 name/font color pairings with three of the background colors (of course, font colors and background colors were always different). Then we divided the resulting 288 name/font-color/background-color combinations into three blocks of 96 combinations so that each target-name/font-color combination occurred only once in each block and so that each name appeared on each of the colored backgrounds at least once in each block. The order of the 96 trials within each block was randomized for each subject separately. In addition, the horizontal location of target name was randomly selected for each trial. Red color font was used for all on-screen instructions and for key labels on the response box.

The color-naming task was followed immediately by an unexpected free recall task in which participants were asked to enter to the computer all target names that they could recall. Participants were given 10 minutes to complete this task.

RESULTS

Recall Data

On average, participants recalled 7.09 of the 24 names presented (32.9%). This is a remarkably low recall rate, given the fact that that each name was presented 12 times and that the names were pretested for high familiarity. The percentage of names recalled did not vary as function of target race or target sex (both *t*s < 1, *n.s.*). Furthermore, the mean adjusted ratio of clustering (ARC) scores (Roenker, Thompson, & Brown, 1971) computed for race and for sex separately were both slightly, though nonsignificantly, negative ($M = -.12$, $t_{(93)} = 1.62$, *n.s.*

and $M = -.10$, $t\ (93) = 1.31$, *n.s.*, respectively), thus providing no evidence of clustering either by race or by sex.

Latencies to Identify Font Colors

Response latencies for identifying the font color constituted the main measure of interest. Latencies shorter than 300 ms[5] (1.1%) and latencies associated with incorrect responses (4.6%) were excluded from the analysis. The remaining latencies were adjusted for effects of serial position, font-color/patch-color combination, and target. Multiple regression (with dummy coding for the last two predictors) was used to compute adjusted latencies. Because preliminary analysis revealed no significant affects involving either participant's sex or target's sex, these variables were dropped from the final analysis. Post-experimental probing confirmed that participants remained naive with respect to the research hypothesis.

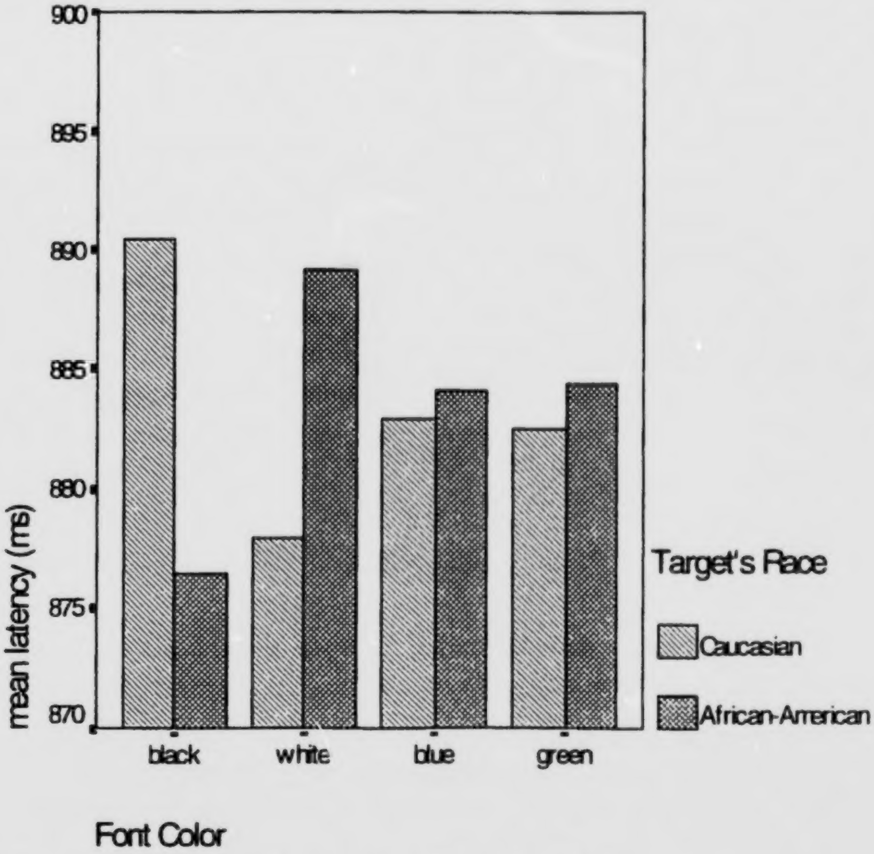

Figure 1 Latencies of naming font color depending on target's race.

As can be seen in Figure 1, the obtained pattern of results was consistent with our predictions. Specifically, recognizing the font color required the least time in trials with the congruent target's race/ font color combinations (the name of a Caucasian target in white letters or the name of an African American target in black letters). Also, recognizing the font color required the most time in trials with the incongruent target's race/ font color combinations (name of Caucasian target in black letters or name of African American target in white letters). A 4 (font color: black vs. white vs. blue vs. green) x 2 (target race: African American vs. Caucasian) mixed model ANOVA performed on response latencies revealed a significant Font Color x Target Race interaction, $F_{(3,\ 91)} = 3.70$, $p < .02$, and no other significant effects. Follow-up tests showed that black font color was recognized faster for African-American than for Caucasian targets, $t_{(93)} = 2.98$, $p < .01$. However, white font color was recognized faster for Caucasian than for African-American targets, $t_{(93)} < 2.46$, $p < .02$. As would be expected, no significant effects of target race were found for either the blue or the green color, both *t*s <1. A corresponding 2x2 mixed model ANOVA performed on response latencies for black and white font colors only, confirmed the predicted specific interaction between the white font vs. black font contrast, on the one hand, and race of the target, on the other, $F_{(1,\ 93)} = 11.13$; $p < .001$. It is worth noting that this critical specific interaction, as well as the omnibus Font Color x Target Race interaction remained significant also when analysis was performed on raw latencies (unadjusted for the effects of serial position, font-color/patch-color combinations, and target). In addition, analysis performed on latencies converted to a logarithmic scale provided similar results.

DISCUSSION

Results obtained with the color-matching task provide strong support for the notion of automatic categorization of persons by race. As predicted, target's race influenced the time required to name the color of the font in which the name was written. Specifically, participants required less time to identify black font when a name of an African-American target was presented in that color than when the name of a Caucasian target was presented. Similarly, participants required less time to identify font color as white when a name of a Caucasian target was presented in that color.

This pattern of results was obtained in spite of the fact that categorizing target persons by race was in no way relevant to the experimental task of naming the font color. Clearly, no overall performance advantage could be gained by adopting the strategy of encoding names of experimental targets in terms of. Our finding that, on average, participants were able to recall only slightly over 30% of the names is consistent with the notion that no conscious effort was made to semantically encode the names in any way. Furthermore the experimental situation lacked any explicit references to either race or to any other type of personal attributes. Thus, the present experiment provides a demonstration of racial categorization as a spontaneous, unintentional process that can occur with no apparent situational provocation.

The present finding of automatic categorization of persons by race should be distinguished from earlier findings (e.g., Devine, 1989; Lepore & Brown, 1997; Wittenbrink, Judd, & Park, 1997) indicating that the race category can be activated automatically by exposure to category labels. In those previous experiments, subliminal exposure to race-category labels (e.g., Blacks, Negroes, Afro-Caribbean) was shown to result in the activation of race-related stereotypes. Arguably, even though category activation was not measured directly in those studies, the obtained pattern of results could not have occurred unless the race-category had been successfully activated. Because the primes were presented subliminally, such activation of race-category would certainly be considered automatic. However, there is a crucial difference between these previous findings indicating that exposure to category labels may result in automatic category activation and the present finding that people may automatically activate and apply racial labels upon exposure to category exemplars (other people). In the first case, category activation occurs as a result of (subliminal) exposure to category labels, in the second case, such activation occurs simply as a result of exposure to category exemplars.

A closer relationship exists between the present experiment and earlier studies that relied on response latencies to assess race-based categorization of person-exemplars (Fazio & Dunton, 1997; Stroessner, 1996; Zarate & Smith, 1990) or activation of (race-related) affective responses to such exemplars (Dasgupta, McGhee, Greenwald, & Banaji,2000; Greenwald, McGhee, & Schwartz, 1998; Ottaway, Hayden, & Oakes, in press). However, with the exception of the experiment by Fazio and Dunton (1997) in which participants were asked to judge similarity between targets that were either matched or mismatched with respect to race, all those studies employed experimental tasks that explicitly called for race-based categorization.

A number of recent models of person perception (e.g., Brewer, 1988; Fiske & Neuberg, 1990; Kunda & Thagard, 1996) make a distinction between top-down, schematic processing involving representations of social categories and bottom-up, highly individualized, attribute-based, processing of social information. Furthermore, all such models assume that the proportional contribution of the first type of processes decreases as target person's familiarity increases. Thus, other things being equal, category labels, perhaps including labels describing race and ethnicity, would play a more prominent role in processing information about a totally unfamiliar target about whom no prior information exists than in processing information about a close friend.

With that distinction in mind, it is relevant to note that our experimental stimuli, well known entertainers, were somewhere between the two extremes of the familiarity continuum. It seems likely that automatic racial categorization effects might have been stronger if less familiar targets were used (e.g., unfamiliar persons identified by race-specific or race-associated attributes, such as physical characteristics, first names, music preferences, speech patterns, etc.). By the same token, it remains an open question if evidence of automatic racial categorization can be found for more familiar targets such as one's social acquaintances, co-workers, class-mates, or friends.

Another unresolved issue concerns the degree to which automatic racial categorization effects occur for targets belonging to specific racial groups. Even though the pattern of results presented in *Figure 1* suggests comparable facilitation/inhibition effects occurring for African American and Caucasian targets, it is important to keep in mind that we used an equal number of African American and Caucasian target persons in our experiment. An unequal distribution, a common occurrence in real life, could have resulted in stronger effects for the minority than for the majority members (cf., Stroessner, 1996; Zarare & Smith, 1990).

In any case, one should not misconstrue the present findings as an indication that categorizing persons according to race occurs universally and "naturally" (see Brewer 1988; Smith & Zarate, 1992, regarding the notion of race as a natural social category) and that such categorization is inevitable. Given the strong affinity between racial categorization on one hand and racial stereotyping and prejudice on the other, this would be an overly pessimistic conclusion. While our results are certainly consistent with the notion, so passionately expressed in our opening paragraph's quotations from Du Bois (1901) and Shipler (1997), of race as a powerful social category, it is nevertheless important to keep in mind the highly arbitrary nature of racial categorization. As noted by Eberhardt and Randall (1997), natural categories are supposed to be invariant and inherently meaningful, yet racial category boundaries are not only highly ambiguous but also tend to shift depending on social, cultural, and historic context.

REFERENCES

Bargh, J. A. & Pietromonaco, P. (1982). Automatic information processing and social perception: The influence of trait information presented outside of conscious awareness on impression formation. *Journal of Personality and Social Psychology, 43,* 437-449.

Bodenhausen, G. V. & Macrae, C. N. (1998). Stereotype activation and inhibition. In R. S. Wyer, Jr. (Ed.), *Stereotype activation and inhibition: Advances in social cognition* (Vol. 11, pp. 1-52). Hillsdale, NJ: Erlbaum.

Brewer, M. B. (1988). A dual process model of impression formation. In T. Srull & R. S. Wyer, Jr. (Eds.), *A dual process model of impression formation: Advances in social cognition* (Vol. 1, pp. 1-36). Hillsdale, NJ: Erlbaum.

Biernat, M. & Vescio, T. K. (1993). Categorization and stereotyping: Effects of group context on memory and social judgment. *Journal of Experimental Social Psychology, 29,* 166-202.

Dasgupta, N., McGhee, D. E., Greenwald, A. G., & Banaji,, M. R. (2000). Automatic preference for White Americans: Eliminating the familiarity explanation. *Journal of Experimental Social Psychology, 36,* 316-328.

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology, 56,* 5-18.

Du Bois, W. E. B. (1903). *The souls of black folk.* Chicago: A. C. McClurg, (Electronic source: Project Bartleby).

Eberhardt, L. E. & Randall. J. L. (1997). The essential notion of race. *Psychological Science, 8,* 198 - 203.

Fazio, R. H. & Dunton, B. C. (1997) Categorization by race: The impact of automatic and controlled components of racial prejudice. *Journal of Experimental Social Psychology, 33,* 451-470.

Fiske, S. T. & Neuberg,S. L. (1990). A continuum of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation. In M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 23, pp. 1-74). New York: Academic Press.

Gilbert, D. & Hixon, J. (1991). The trouble of thinking: Activation and application of stereotypic beliefs. *Journal of Personality and Social Psychology, 60,* 509—517.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association task. *Journal of Personality and Social Psychology, 74,* 1464-1480.

Hewstone, M., Hantzi, A. & Johnston, L. (1991). Social categorization and person memory: The pervasiveness of race as an organizing principle. *European Journal of Social Psychology, 21,* 517-528.

Hamilton, D. L. & Sherman, S. J. (1994). Stereotypes. In R. S. Wyer, Jr. & T. K. Srull (Eds.), *Handbook of social cognition* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Klauer, K. C. & Wegener, I. (1998). Unraveling social categorization in the "Who Said What?" paradigm . *Journal of Personality and Social Psychology, 75,* 1155-1178.

Kunda and Thagard (1996). Forming impressions from stereotypes, traits, and behaviors: A parallel-constraint-satisfaction theory. *Psychological Review, 103,* 284-308.

Lepore, L. & Brown, R. (1997). Category and stereotype activation: Is prejudice inevitable? *Journal of Personality and Social Psychology, 72,* 275-287. MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin, 109,* 163-203.

Ottaway, S. A., Hayden, D. C., & Oakes, M. A. (in press). Implicit attitudes and racism: The effect of word familiarity and frequency in the Implicit Association Test. *Social Cognition.*

Roenker, D. L., Thompson, C. P., & Brown, S. C. (1971). Comparison of measures for the estimation of clustering in free recall. *Psychological Bulletin, 76,* 45-48.

Shipler, D. K. (1997). *A country of strangers: Blacks and whites in America.* New York : Knopf

Smith, E. R. & Zarate, M. A. (1992). Exemplar-based model of social judgment. *Psychological Review, 99,* 3-21.

Stangor, C., Lynch, L., Duan, C. & Glass, B. (1992). Categorization of individuals on the basis of multiple social features. *Journal of Personality and Social Psychology, 62,* 207-218.

Stroessner, S. (1996). Social categorization by race or sex: Effects of perceived non-normalcy on response times. *Social Cognition, 14,* 247-276.

Spencer, S. J., Fein, S., Wolfe, C. T., Fong, C., Dunn, M. (1998). *Personality and Social Psychology Bulletin, 24,* 1139-1152.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology,18,* 643-662.

Taylor, S. E., Fiske,S. T., Etcoff, N. L. & Ruderman, A. (1978). Categorical and contextual bases of person memory and stereotyping. *Journal of Personality and Social Psychology, 36,* 778-793.

Wittenbrink, B., Judd, C. M. & Park, B. (1997). Evidence for racial prejudice at the implicit level and its relationship with questionnaire measures. *Journal of Personality and Social Psychology, 72,* 262—274.

Zarate, M. A., & Smith, E. R. (1990). Person categorization and stereotyping. *Social Cognition, 8,* 161-185.

Author Note Preliminary results of this experiment were presented at the 1998 Annual Meeting of the Association for Experimental Social Psychology in Lexington, Kentucky. Michael Motes is now at Texas Christian University. We thank Christopher Martin for helpful comments on an earlier draft of this manuscript.

Footnotes

---

[1] The question of whether people automatically engage in racial categorization (categorize persons according to race) should be distinguished from a related question of whether the race category (and associated racial stereotypes) can be activated automatically as a result of exposure to category labels (e.g., Devine, 1989; Lepore & Brown, 1997; Wittenbrink, Judd, & Park, 1997). We will return to this distinction in the Discussion.

[2] In fact, those studies provide a good illustration of the need to distinguish between categorization and stereotype activation. Specifically, results indicated that, under specific conditions (high cognitive load and positive feedback), activation of stereotypical content did not occur. Without a direct measure of categorization, it is impossible to know whether those conditions prevented participants from categorizing targets as African Americans or whether negative stereotypical content was not activated, even though categorization did occur.

[3] An auxiliary analysis revealed that, with non-Caucasian participants included, all reported results remained significant.

[4] The names were selected from a larger pool based on familiarity ratings obtained from a separate sample of undergraduates.

[5] The value of 300 ms was chosen arbitrarily after examining the distribution.

An Interview With Kimmo Lehtonen: Music Therapy With Adolescents

Kimmo Lehtonen
University of Turku, Finland

(interviewed on behalf of NAJP by)

Michael F. Shaughnessy
Eastern New Mexico University

Music therapy is becoming increasingly recognized as a form of treatment with adolescents. In this interview, Dr. Kimmo Lehtonen discusses the current state of the art of music therapy as it relates to adolescents and gives examples derived from his own practice of music therapy in Finland and Scandinavia.

NAJP: Please describe your current work with adolescents.

KL: I currently work near Turku, Finland at a Family Rehabilitation Center. They have a strong family therapy orientation or ideology and family therapy and music therapy seems to complement each other very well. The adolescents stay there from 6 months to 6 years because it is a residential facility.

It is entirely voluntary. One does not have to come and participate. Often people want to come and simply make music. They have heard about the fact that people are making music together. They want a feeling of camaraderie and it is not a good word in Finnish—because it means Soviet Union or Russian word comrade which is a military term, but they want to belong and enjoy the feeling of working together.

If you are making good music with a group you have to listen, you have to play your part and you have to have respect for others. The ultimate goal of the Family Rehabilitation Center is to alleviate family symptoms.

NAJP: Why is music therapy so successful with adolescents?

Author info: Correspondence should be sent to: Dr. Michael Shaughnessy, Eastern New Mexico University, Portales, NM 88130.

North American Journal of Psychology, 2002, Vol. 4, No. 1, 13-20.
© NAJP

KL: I think that music comes from and touches inner feelings of human beings, and in my situation it is often so that adolescents cannot express themselves through words. They have already found music to be a self-healing element and they have already found a love of music, and it is already a big part of the adolescent sub-culture. I think music is a type of emotional play where the adolescent is telling about his/her feelings in a safe environment. So they are able to relate to music better, and using music is a way to "step in to" their culture. They are able to indicate why they like it and how it relates to their lives and their problems.

I think many of these students are somewhat anti-social. They don't like school, their teachers, their parents, and this is a result of their coming from broken homes. They just don't trust adults. I don't mean they are clinically anti-social, but anti-social in the sense of their feeling toward "societal others" or significant others in their lives.

Almost everybody I see comes from a broken home. It is an exception that I see kids from an intact home. A broken home is not the worst thing or the only thing, there are other things that come together, but it is easy to point to the "broken home" as a cause. The broken home is often the "tip of the iceberg" to use Freud's terminology. There are often underlying causes in drug and alcohol abuse, emotional abuse.

I think sometimes their parents simply can't function as parents. They don't have the skills, and they have negative backgrounds and feelings toward their own parents.

Many adolescents try to "escape the family" too early by joining gangs and beginning sexual relations too early. The gang leader is often a twisted father figure. They begin sexual intimacy before they have the skills for emotional intimacy. Girls try to get emotional love when they didn't get it from their parents and they try to get respect by having a boyfriend who has respect. Some girls try to get their self-esteem validated by having a boyfriend who is respected, thinking this will contribute to their self-worth.

NAJP: How can we best train music therapists to work with adolescents?

KL: The best teachers are sometimes those adolescent kids. If one wants to learn to swim, one has to go near the water. One needs to be conscious of one's own development and it is important to have worked through one's own problems. In a sense, one is going to be young again when you work with adolescents. One has to remember not to lose one's adulthood either. Those kids need a trustworthy adult, not an overage adolescent.

Good supervision is also important. The hardest patients are probably the ones who have the same kinds of problems that you had. The relationship you have with adolescents is also important.

Therapists need a good "tool box." You need some training in projectives, improvisation songs, and you have to be aware of the youth culture and trends and be prepared to improvise. Sometimes adolescents can learn that they can be good in at least one thing—maybe learn to play by ear or pick up chords or follow the beat.

NAJP: What are you currently researching?

KL: My last project involved music thought to be important by Finnish psychiatric patients. We sent a questionnaire to all the hospitals in Finland and asked patients to name the 10 most important pieces of music in their lives. We found out that these pieces were found to reflect important developmental milestones. Adolescence was seen to be a critical time. The pieces typically were found in their teenage years. Interestingly, 23 of 25 had been written in minor keys. They were pervasively sad and told about failure, about losing a love. I noted a lot of Jungian archetypes in these songs, such as crossing a bridge, crossing the water, shadows, and many dealt with this Jungian concept of "anima." The male artists were singing songs to women and about women. The Eagles' "Hotel California" was noted to hold many symbolic themes—the woman with a candle, leading the way.

"Yesterday," by the Beatles, with "shadows hanging over me," was another deeply felt song with themes of loss. "Bridge Over Troubled Water" also reflected the transitional nature of many patients having to take giant steps, leaving behind a loved one , a marriage, one's youth. This may also reflect the generalization that Finlanders are melancholic, or it may reflect the transitional problems associated with loss.

NAJP: Why is adolescence such a difficult time?

KL: One has to develop as a human being, and it is one's *last chance* to form oneself as an adult or before one becomes an adult. It is one's last chance to work through those things in adolescence. I am currently writing an article on this aspect of adolescence. For some teenagers, adolescence is also a first change - the first date, the first kiss, the first sexual encounter. It is also a last chance for many teenagers to experience the freedom of youth and adolescence before they have to accept adult responsibilities. Adolescence, as we all know, is a time of transition and the identity crisis. We try to help them through this crisis in music therapy. I think it is much more difficult to work through these issues as an adult.

Adolescence is a dynamic time when one is more flexible. It is a time of seeking one's self and seeking one's favorite things and seeking other people. I think I enjoy working with adolescents since it is a dynamic time. Adults are too rigid and set in their ways.

Adolescence may also be thought to be one's last chance for happiness before adulthood. It also may be regarded as the last chance to form one's self-esteem and self-worth. This is the "last chance" to resolve their perceived problems.

It is almost as if adolescence is their last chance before they cross that bridge into adulthood. I think that the bridge motif seen in many songs is quite important. There are many adolescents who are standing on this bridge looking at this water, trying to cross this bridge. For some it is a time of optimism, for others despair.

NAJP: What is currently happening with adolescents in Scandinavia?

KL: I think there is a greater polarization between A class and B class. The A class has more education, motivation and drive toward success. The B class is often termed "losers" (Luuser). This reflects what's happening in society—we see these kids drop out of school, society work. It's very much a self-fulfilling prophecy of expectations. Others do not expect them to do well and they don't expect to do well.

Teenage pregnancy is somewhat down due to fear of AIDS and STD's. In the last 5 years, hashish, amphetamines, marijuana and heroin use has increased. It's very obvious at least to me. Sniffing and snuffing has increased and other substitutes have increased. Mixing alcohol and hashish has also increased. Drugs are coming from the former Soviet Union into Scandinavia and it is difficult to control this.

NAJP: Why do so many adolescents consider suicide?

KL: They are really more destructive than suicidal—they are more self destructive than intent on killing themselves. Many kids are depressive and have serious troubles and during their development life has been a cold lover—it has asked a lot without giving anything. Everything they have learned at school is that they are not able to do anything. In therapy I say "Let's sing," and the adolescent says "I can't sing." I say "let's play," and they say "I can't play," and I say "let's talk about it," and they say "I can't talk about it." Often I put them behind drums and they can begin to communicate—everybody has some type of rhythm.

Often it becomes obvious that the adolescent has never learned to listen. They don't listen to the other players or the rhythm or the beat. They have never learned to listen because they have always been criticized. In music therapy you have to listen to intentions, motives and meanings.

Often therapists also need to listen to what the client is playing and how they are playing. The style of playing tells a lot about what they are feeling. Let me give a few examples to explain. Once I had a little quiet, introverted, shy girl, who was playing a piano so quietly that I could hardly hear her playing! This reflected her personality and her style of interaction.

On the other hand, I had an aggressive, violent, destructive boy who started to play a very loud, screaming, aggressive guitar with feedback and reverb and with such volume that my ears were humming for two days after that!

I have had music therapy groups where some of the players were trying to play too fast, and the others found it very hard to either follow, or to keep up. We have to work together, cooperate and communicate, and be considerate of others, and music therapy helps us to do just that. It also helps us to express ourselves, deal with our feelings and emotions and inner psychic trauma and issues in an important way!

NAJP: How can music therapy help emotionally disturbed adolescents?

KL: Music can reflect the inner feelings and activate feelings you never knew you had. You can reflect your thoughts and ideas and mental pictures using music. I had one adolescent who liked Kurt Cobain and his destiny—and this adolescent was using Cobain to talk about himself—his own ideas about suicide.

NAJP: Why is music therapy successful when traditional counseling "talk therapy" is not?

KL: The main thing is that you don't have to tell anything, you simply express yourself. Many kids have met "healers" such as psychiatrists, and have always been asked the same questions. They feel the questions are stupid and they are tired of listening to their advice.

So I often use a paradox saying you don't have to say or tell anything, and I don't even read the case histories—it's more important to form your own picture or perspective of a kid. Music also has a more direct connection to one's unconscious mind. Pieces you like the most tell a lot about your unconscious mind. Often I suggest that the music is telling about your life. The idols you have reveal much about your own identification. One adolescent liked Jim Morrison, Jimi Hendrix and Kurt Cobain. I pointed out that they were all dead. He had a record entitled "I Woke Up One Morning and Found Myself Dead." It was recorded by Jim Morrison, Jimi Hendrix and Johnny Winter.

NAJP: Are there any studies comparing the effectiveness of music therapy with other kinds of therapy?

KL: I don't think that there are any longitudinal or large-scale studies comparing the effectiveness of music therapy with other kinds of therapy. I think that we need these kinds of comparative evaluations, but such large-scale investigations, at least to my knowledge, do not exist. It would be wonderful if we had the funds and funding to conduct such investigations and examine a number of different variables, so as to determine the most relevant factors and salient aspects of music therapy. Perhaps some funding agency or organization will read this and support this type of study. But, in general, case studies are the avenue of research and investigation.

NAJP: What question have I neglected to ask?

KL: Perhaps a question about the necessary personality of the music therapist. He/she must give the client space, choose the correct music and become a warm parental figure that the client has not had. I never adopt the professional role—I give space too. I never try to give a life model saying you should do this, do that. I remember one guy who was stealing cars. I said "You should go on stealing cars as long as you have a need to steal cars. The cars are only a substitute for whatever you do not have. Other kids have fathers—perhaps you steal cars because you do not have a father."

NAJP: Have you ever written about the idea that music has a personality?

KL: Yes, I have one article in Finnish I need to translate. I think that the main question is whether or not music is a living organism. I am using Karl Popper's idea about the 3 worlds—I think that music is a living organism from the 3rd world. It functions as a subject—when I play a record, it forms mental pictures, associations and memories are resurrected of the lost past. Music is like a subject which causes many kinds of changes in both the therapist and the patient or client. The third world theme, is taken from the great philosopher, Karl Popper, who is talking about three worlds. The first world, is the world of physical objects. The second world contains consciousness and mental stages, and the final world, number three, contains certain thoughts, theories, works of art , pieces of music and so on.

According to Popper, this world is man made, but only relatively man made because it is relatively autonomic and independent and because it has its effects on human beings. For instance in the music therapy process, a certain piece of classical music can bring its independent effect on the process. Let me give an example. I was once working with an 8 year old girl and I used the Symphony Fantastique (by Hector Berlioz) and suddenly this girl started to tell me about the contents of the piece (it has a program written by Berlioz). The girl's story was quite close to the original program. I have written an article in Finnish entitled "Is Music a Subject," in which I asked this question: "Can a piece of music act as a subject, having its effects on us in different times and in different places and different cultures?"

Can it have its individual unique meanings that start a listener's different psychological processes in different places? Can this be described so that, for instance, Berlioz has made an independent subject (of the third world) which acts a little bit like a "living organism" and takes its part in the music therapy process. The intention of this subject is to touch people in its own personal way. This is a long and perhaps too complicated story to explain with a few words. It needs an entire article to explain.

A theory can be like a subject too. It inspires researchers and research over and over again during decades and centuries. It has its effects everywhere. In a way, we can say that it is living with various researchers everywhere and has a life of it's own which is not dependent on its creator.

REFERENCES

Lehtonen, K. (1992). Music therapy in Finland. *Nordic Journal of Music Therapy 1*, 30-32.

Lehtonen, K. (1993) Finnish music therapy in a nutshell. *Finnish Journal of Music Therapy*. Special English Issue. 1-2, 5- 13.

Lehtonen, K. (1994) Gibt es Entsprechung zwishen Structuren von Musik und denen der Psyche? *Musikterapeutische Umshau. Forschung und Praxis der Musiktherapie. 1*, 9-24.

Lehtonen, K (1994) Relacio entre la Musica y la Pscioterapia. In Hensky de la Gaiza, V. (Ed) La education Musical Frente al Futuro. Enfogues interdisciplinaros desde la Filosofia, La Sociologia, la Antropologia, la Picologia, la Pedagogia y la Terapia. 129-137. Buenos Aires: Editorial Guadalupe.

Lehtonen, K (1995). Musicotherapia. Relacio entre Musica y Psicoterapia. *Revista Internacional Latino America de Musicotherapia* 2, 21- 44.

Lehtonen, K. (2001) Music in Life-Space: A Lewin–like viewpoint to the Meaning of Music and Music Therapy. *Canadian Journal of Music therapy.* 7, (1), 19-38.

The Relationship of Verbal-Nonverbal Incongruence to Communication Mismatches in Married Couples

Amy Van Buren
Sacred Heart University

Communication accuracy refers to whether a message sent by a sender is perceived by the receiver to have the same emotional meaning intended by the sender. Previous research using marital dyads suggests that receivers sometimes receive the emotional meaning in senders' statements differently than senders intend. The present study was conducted to test the possibility that one reason such misunderstandings occur is that senders may convey emotional messages differently than they intend. Twenty-four married couples carried on a ten-minute videotaped free interaction during which they rated the emotional meaning in each others' statements. Results indicated that senders conveyed messages that were both more negative and more positive than they intended. As predicted, emotional mismatches, in which sender and receiver disagreed on how the sender was coming across, were associated with sender verbal-nonverbal incongruence. However, it was also found that matches, in which sender and receiver agreed on how the sender was coming across, were associated with incongruence. Contrary to prediction, when senders were incongruent, their verbal and not their nonverbal behavior correlated significantly with the impact on the receiver. The finding that receivers' impact ratings correlated more with senders' verbal than nonverbal behavior contradicts results from previous laboratory-based studies on the resolution of discrepant verbal and nonverbal signals. It suggests that nonverbal behavior may be best studied and interpreted in terms of its verbal context.

It is generally agreed among theorists and clinicians who study dyadic communication that, except in some special circumstances dictated by custom, one of the requirements for healthy communication is that the message intended by the sender is the same as the message received by the listener (Duke & Nowicki, 1982; Kiesler, 1996). In particular, it is especially important that the emotional meaning consciously intended by the sender is the same as the emotional meaning received by the listening party. Noller (1984, 1992) has termed this skill "communication

---

Author info: Correspondence should be sent to: Dr. Amy Van Buren, Psychology Dept., Sacred Heart University, 5151 Park Ave., Fairfield, CT 06432.

North American Journal of Psychology, 2002, Vol. 4, No. 1, 21-36.
© NAJP

accuracy." For example, in a discussion with his wife about her mother, a husband might utter the sentence, "I like your mother" and may intend to convey a sense of liking. If his wife hears him as positive, an accurate communication has occurred. However, if his wife hears him as defensive or argumentative, a mismatch has occurred in which the husband's intended communication is different from its impact on his wife. The ensuing statements would likely be very different from and more negative than those that would follow if the wife received the husband's statement in the way he intended.

Accuracy of communication has long been recognized by interpersonal and communications theorists as an important element of interpersonal interactions. For example, in discussing the complexities of human interaction, Satir (1964, 1976) suggested that repeated emotional misunderstandings form the basis of unhealthy interactions and occur because senders and receivers often misinterpret each other. According to Satir, such misunderstandings occur when senders and receivers "fail to check out meaning intended with meaning received" (1964, p.100).

While in theory accuracy in communication is certainly a healthy goal, research using marital dyads suggests that is not a goal that is always easily attained in naturally occurring dyadic interactions (e.g., Gottman, Notarius, Markman, Yoppi & Rubin, 1976: Noller, Feeney, Bonnell & Callan, 1994; Noller & Guthrie, 1989; Schacter & O'Leary, 1985; Sillars, Folwell, Hill & Maki, 1994; Vangelisti, 1994). In studying communication accuracy in couples, for example, Gottman and his colleagues have used a device called a "Talk Table," which makes use of couples' own perceptions of their conversations (Gottman, et al., 1976). The Talk Table operates by recording both the speaker's intent and the resulting emotional impact on the listener. Spouses' intent and impact ratings are recorded by spouses themselves on a five point scale ranging from "superpositive" to "supernegative." Using the Talk Table to record communication mismatches, Gottman found that the communication of both distressed and non-distressed couples contained mismatches, and that distressed couples in particular tended to perceive each other's messages as more negative than intended (Gottman, et al. 1976).

Using a similar methodology, Schachter and O'Leary (1985) found that the communication of both distressed and non-distressed couples contained mismatches in which the sender's impact was perceived as less positive than intended. Taken together, Gottman et al.'s and Schachter and O'Leary's findings suggest that emotional misunderstandings frequently occur in the communication of married couples, with both groups tending to perceive the sender's message as less positive than intended. However, an important question, which has not been adequately addressed in the marital literature, is whether it is the sender or the receiver (or both) who is contributing to the error. The assumption made by many researchers is that receiving spouses may erroneously interpret their partners' messages as more negative than the messages are intended (Filsinger & Wilson, 1983; Schaap & Jansen-Nawas, 1987). However, as Noller (1984) has noted, it is also possible that spouses may *send* their messages more negatively than they realize.

Noller (1980) found evidence for this when she conducted a study using a variation of the Marital Communication Scale (Kahn, 1970). The MCS consists of several sentences which can convey positive, negative, or neutral meanings, depending on the nonverbal signals that accompany the words. Sending spouses in the study were instructed to send messages to each other in one of three ways, and receiving spouses were instructed to guess from a list of three alternatives which ideas senders were trying to convey. Senders' messages were videotaped and later shown to a group of judges who were also instructed to guess the meaning conveyed by sending spouses. Sending spouses were judged as having made an error if the receiving spouse plus more than one-third of the judges heard the message differently than the sender intended it. It was found that distressed senders, particularly husbands, made significantly more errors in communicating than non-distressed senders. Errors were most likely to occur when a sender attempted to send a positive message, but actually sent a message judged by the receiving spouse and judges to be neutral or negative.

While Noller's results point to the possibility that emotional misunderstandings may in part be due to sender error, the question still remains as to what kind of error senders are making. An answer may be found in the interpersonal formulations of Kiesler (1979, 1986, 1996). Kiesler has built upon Sullivan's interpersonal theory (1953), which posits a central role for interpersonal communication in personality development and maintenance. According to Kiesler, it is possible for senders to be unaware of the emotional messages they are communicating but which nevertheless have an emotional impact on receivers. This can happen because communication takes place in two channels: one verbal, the other nonverbal. While the verbal channel consists of words, the nonverbal channel consists of behaviors that accompany the words. Nonverbal communication consists of several behaviors, including facial expression, paralanguage (voice tone, pitch, loudness), kinesics (posture and gesture), and proxemics (utilization of space and distance) (Burgoon, Buller & Woodall, 1996; De Paulo, 1992; Hickson & Stacks, 1985; Knapp & Hall, 1997; Riggio, 1992).

According to interpersonal theory, senders sometimes have needs whose direct expression might elicit unpleasant feedback (Beier & Young, 1984). Senders in such situations often attempt to convey their needs while simultaneously disavowing responsibility for them (Beier & Young, 1984). They may do this by consciously conveying one message in the verbal channel while simultaneously and unconsciously transmitting another in the nonverbal channel. When this happens, it is the nonverbal message that is the more powerful of the two and that leaves an emotional impact on the receiver (Capella & Palmer, 1989; Van Denburg, Schmidt & Kiesler 1992). The result is that the receiver perceives a different emotional meaning than the one consciously intended (and verbally sent) by the sender and an emotional misunderstanding occurs between the members of the dyad.

The present study was designed to investigate whether interpersonal assumptions regarding the determinants of sender incongruence could be used to explain emotional misunderstandings that occur in dyadic communication. Communication accuracy was measured using a variation of the Talk Table methodology developed by Gottman and his colleagues. The Talk Table allows for a quantitative assessment of communication accuracy by comparing the sender's intent with the resulting impact on the receiver. Non-clinic married couples were used, since such a sample contains naturally occurring dyads in which it has been shown that communicative misunderstandings occasionally occur (Gottman, et al., 1976; Schachter & O'Leary, 1985).

The present study tested two predictions. The first was that mismatches, in which the message sent by the sender was different than the message perceived by the receiver, would be characterized by sender verbal-nonverbal incongruence. Conversely, it was predicted that when sender and receiver agreed on the emotional tone of the sender's message (an instance referred to as a communication "match") the sender's behavior would not contain verbal-nonverbal incongruence. The second prediction held that, when senders were incongruent between verbal and nonverbal signals, receiver impact ratings would correlate with senders' nonverbal behavior (specifically, facial expression and voice tone) and not their verbal language.

METHOD

Participants

Participants were 24 Caucasian married couples from Westchester County (New York), New York City, and New Jersey. Couples were selected from lists given the experimenter by ministers and rabbis in each area. Approximately 94% of couples contacted agreed to participate.

Length of marriage ranged from 2 - 52 years (*M* = 14 years, *SD* = 13 years). Nineteen out of the 24 couples were in their first marriage. In the remaining five, either one or both of the spouses had been previously married.

Raters

Three raters were used to code videotapes, audiotapes and transcripts. All were in their late 20's and two were enrolled in graduate psychology programs. The third had a B.A. in psychology. Two female raters coded the entire sample, and the third, a male, was used for purposes of a reliability check. All three were trained.

Measures

Variation of Gottman et al.'s (1976) Talk Table. Communication accuracy was measured using a variation of Gottman et al.'s Talk Table. As used in the present study, the talk table is a pencil and paper self report measure designed to allow couples to record emotional meaning intended (intent) and emotional meaning received (impact). The talk table in the present study contained three words to describe emotional meaning: "positive," "neutral," and "negative." The function of the talk table is to record potential differences between spouses' intents and impacts.

The Talk Table was used in the following manner. Spouses were instructed to carry on a conversation with each other. After each minute of interacting, spouses circled the response that best fit 1) their intended impact during the past minute and 2) the impact of their spouse's communication during the past minute. Items directed toward senders were preceded by this stem:

"During the past minute I intended to come across as...." Items directed toward receivers were preceded by the stem, "During the past minute, my spouse came across as...." Each stem was followed by the choices "positive," "neutral, " or "negative."

Procedure

All couples were contacted by phone during which time the experimenter explained the purpose of the study and outlined the details of participation. If both husband and wife agreed to participate, a date was set for an interview. All couples were interviewed and videotaped in their homes. During the interview, the experimenter explained that she was interested in the way in which married couples talk about issues on which they disagree. Couples were asked to come up with an issue on which they disagreed and which they would be willing to discuss in front of a video camera. Once both spouses and the experimenter agreed on a topic for discussion, the experimenter explained the talk table procedure. The meaning of each talk table term was explained. Spouses were to endorse "positive" if the emotional tone of the conversation was upbeat and satisfying, "negative" if the conversation was argumentative or upsetting, and "neutral" if there wasn't a particular emotion being displayed either way. Spouses were not given any instructions to focus on a particular nonverbal channel; rather, they were instructed to focus on each other's overall behavior.

All couples were videotaped sitting on a sofa. Each spouse was given a form on which to record intent while speaking and impact while listening. Couples were then instructed to discuss a problem or issue about which they disagreed and which had been previously agreed upon. After each 60 seconds of interaction, a previously made tape-recorded interaction instructed each spouse to record 1) his/her intent for the previous 60 seconds and 2) the impact of his/her spouse's communication during the previous 60 seconds. A total of 20 ratings were made by each spouse, 10 for him/herself and 10 for his/her spouse. Spouses' intents and impacts were recorded by couples on rating sheets that contained the talk table rating scale.

After all talk table data were collected for all 24 couples, the recording sheets were scanned for mismatches. Mismatches consisted of all 60-second interaction sequences in which the speaker's intent differed from his/her impact on the listener. For each mismatch, three pieces of data were assembled in order to separate the nonverbal from the verbal element of the sender's behavior. The three pieces of data were 1) the verbal dialogue (written transcript), 2) a videotape of the speaker's facial expression and 3) an audio recording of the interaction. The purpose of separating verbal from nonverbal elements of senders' behavior was to determine whether mismatches were associated with particular kinds of sender verbal-nonverbal incongruence. There were two kinds of sender incongruence possible. The first was verbal-facial incongruence, in which the sender's facial expression conflicted with his/her verbal language (for example, the sender's face showed negativity while his/her words were positive). The second kind of sender incongruence was verbal-vocal, in which the sender's tone of voice conflicted with his/her verbal language (for example, the sender's tone of voice was neutral while his/her words conveyed negativity).

Using the same talk table data scales on which spouses had recorded intents and impacts, raters then coded speaker's verbal, facial, and vocal behavior during each 60-second mismatch. Verbal behavior was rated using a written transcript and facial expression was rated using the video of the speaker's face. In order to rate speaker's vocal behavior, raters listened to an audio recording of each 60-second mismatch and were instructed to ignore the content of what was being said. Raters were given specific instructions for rating the kind of emotion (positive, negative, or neutral) present in each channel. Training criteria for the nonverbal channels of facial expression and vocalics were adapted from a nonverbal rating scheme developed by Mehrabian (1981). Verbal training criteria were based on the verbal "social skills" section of Gottman, Notarius, Gonso & Markman's *A Couple's Guide to Communication* (1976).

Rater Training

Three raters were used, two of whom rated the entire sample. The third rated 25% of the sample as a reliability check. Rater training involved initially reading and discussing criteria for positive, neutral, and negative behaviors. All three raters then made practice codings on pilot tapes and transcripts that were not used in the actual study. Raters were instructed to focus on the behavior or words of the spouse in question and, based on training criteria, to select the emotion (positive, negative, or neutral) most prevalent during that 60 second sequence. The three raters made their judgments separately, but compared and discussed instances in which they disagreed. In each case of disagreement, reasons for disagreement were discussed and a new coding criterion was established. After training, correlations of rater agreement for each channel of behavior were computed. Pearson correlations computed between practice codings made by the two main raters were as follows: verbal channel (words) $r = .83$; visual channel $r = .93$; vocal channel $r = .82$. This was taken as an indication of sufficient inter-rater reliability and coding of the actual sample was begun. In each instance in which the two main raters disagreed on a score to be given a particular behavior, the third rater was used to code the behavior in question, and the final rating was based on a 2/3 majority. Instances in which the two main raters disagreed and the third rater was used to resolve the deadlock occurred approximately 10% of the time.

RESULTS

Mismatches were separated into two groups: 1) positive mismatches, in which senders rated their overall behavior as negative or neutral while receivers perceived them as positive and 2) negative mismatches, in which senders rated their overall behavior as neutral or positive while receivers perceived them as negative. Analyses were conducted separately for positive and negative mismatches.

The first prediction was that the communication mismatches that occur in the conversation of married couples would be associated with sender verbal-nonverbal incongruence. Pearson correlations were conducted between the number of positive or negative mismatches per couple that were associated with sender verbal-facial or verbal-vocal incongruence. As predicted, positive mismatches were significantly correlated with both verbal-facial ($r = .52$, $p<.01$) and verbal-vocal ($r = .38$. $p<.025$) incongruence. Likewise, as predicted, negative mismatches were also significantly associated with sender verbal-facial ($r = .84$, $p<.0005$) and verbal-vocal ($r = .74$, $p<.005$) incongruence. A Fisher's Z test for the difference between correlation coefficients indicated that there were no significant differences between the correlations between negative and positive mismatches with verbal-facial incongruence ($Z = 1.71$) or between the correlations between negative and positive mismatches with verbal-vocal incongruence ($Z = 1.46$). In percentage terms, the total number of both positive and negative mismatches associated with incongruence was 60%.

Data from positive and negative *matches* (in which sender and receiver agreed on how the sender was coming across) were also analyzed for incongruence in order to ascertain whether sender incongruence was in fact specific to mismatches. Findings indicated that, contrary to prediction, matches were also significantly associated with both verbal-facial and verbal-vocal incongruence. Positive matches were significantly associated with sender verbal-facial ($r = .74$, $p<.005$ and verbal-vocal ($r = .71$, $p<.0005$) incongruence. Likewise, negative matches were also significantly associated with verbal-facial ($r = .73$, $p<.005$ and verbal-vocal ($r = .85$, $p<.005$) incongruence. In terms of percentage, 62% of the total number of matches were associated with either verbal-facial or verbal-vocal incongruence.

The second prediction was that in communication mismatches in which there was sender incongruence, the sender's nonverbal behavior would correlate significantly with the emotional impact on the receiver while the sender's words would not. For positive and negative mismatches, correlations were computed between the number of times per couple that the receiver perceived the sender as positive or negative and the number of times the sender's face, voice, or words agreed with the receiver's score.

In all cases, it was found that, contrary to prediction, senders' words correlated significantly with receiver impact, while senders' facial expressions and voice tone did not. For positive mismatches with verbal-vocal incongruence, voice tone was not significantly associated with receiver impact ($r = -.14$), while words correlated significantly ($r = .58$,

$p<.05$) with receiver impact. Similarly, for positive mismatches with verbal-vocal incongruence, senders' words were significantly associated with receiver impact ($r = .69$, $p<.01$) while voice tone was not ($r = .21$). For negative mismatches with verbal-facial incongruence, senders' words were significantly correlated with receiver impact ($r = .75$, $p<.01$),while facial expression was not ($r = .35$). However, a Fisher's Z test for the difference between correlation coefficients indicated that the verbal (words) and facial correlations were not significantly different from each other ($Z = 1.26$, $p>.05$). For negative mismatches with verbal-vocal incongruence, senders' words were significantly associated with receiver impact ($r = .53$, $p<.05$) while sender voice tone was not ($r = .38$).

Data from positive and negative matches also were analyzed to determine whether any one specific channel correlated significantly with receiver impact when the sender was incongruent. Again, in cases of both positive and negative matches with sender incongruence, senders' words were significantly associated with receiver impact while senders' nonverbal behavior was not. For positive matches with verbal-facial incongruence, senders verbal behavior correlated significantly with receiver impact ($r = .89$, $p<.0005$), while senders' facial expression did not ($r = .04$). For positive matches with verbal-vocal incongruence, senders' verbal behavior correlated significantly with receiver impact ($r = .96$, $p<.0005$) while sender vocals did not ($r = .28$). For negative matches with verbal facial incongruence, sender verbal behavior correlated significantly with receiver impact ($r = .94$, $p< .0005$) while sender facial expression did not ($r = .00$). For negative matches with sender verbal-vocal incongruence, sender verbal behavior correlated significantly with receiver impact ($r = 1.00$, $p<.0005$), while sender vocals did not ($r = .02$) .

DISCUSSION

The findings in this study bring into question two predictions based on interpersonal theory. First, while sender verbal-nonverbal incongruence was found to occur during misunderstandings, as predicted, it also was found to occur when spouses were communicating accurately, a finding not anticipated. Second, contrary to prediction, it was found that when senders communicated incongruently, receivers' impact ratings correlated significantly with senders' verbal language rather than senders' nonverbal behavior.

The finding that senders were incongruent regardless of whether there was a match or a mismatch suggests that sender verbal-nonverbal incongruence may be a much more naturally occurring phenomenon in communication than interpersonal theorists would expect, and that couples can be incongruent but still communicate accurately. This parallels Noller and Gallois' (1988) finding that even spouses who are highly accurate communicators sometimes communicate incongruently, for example by sending a negative verbal message with a smile. In fact, it may not be sender incongruence per se that contributes to a mismatch, but whether the incongruent sender is aware of the channel carrying the message to which the receiver responds. For example, a husband, in uttering a negative sentence with neutral face and voice tone may nevertheless realize that it is his words that carry the emotional weight of his utterance. If his wife then responds to the negative meaning contained in his words (but not in his face and voice), he is not likely to be surprised by his response to her, and an accurate, although negative, communication will have taken place between them. Support for this possibility has been found in a study on communication awareness among married couples, in which high marital adjustment husbands were more aware than low marital adjustment husbands of how accurately they encoded messages (Noller & Venardos, 1986).

It thus follows that those senders who are more aware of what channel receivers are paying attention to are those who have more accurate communication with their spouses. This would explain how it was possible for couples in this study to communicate incongruently but still communicate accurately. One way to address this in future research would be to ask senders which channel they thought the receiver would respond to in making the impact rating.

The finding that receivers' ratings correlated more with senders' verbal language than nonverbal cues also contradicts an assumption from interpersonal theory that the nonverbal channel is the one that carries the emotional message when verbal and nonverbal signals conflict. Findings indicated that, in all cases of sender verbal-nonverbal incongruence except that of negative mismatches, senders' words correlated significantly with receiver impact while senders' facial expression and voice tone did not. Where negative mismatches were concerned, while senders' verbal behavior was significantly associated with receiver impact and their nonverbal was not, there were no significant differences between the verbal and nonverbal correlations themselves. This suggests that senders' nonverbal behavior may also have contributed to the impact on the receiver when a negative mismatch occurred. Negativity also may have been carried through senders' faces and voices as well as through their words.

The above findings would seem to bring into question assumptions made by interpersonal theorists regarding the relative importance of nonverbal and verbal sources of emotion. However, it is important, before reaching this conclusion, to first examine the validity of the methodological practices used in this study that also might have contributed to the results. Specifically, there are several possible methodological reasons why senders' nonverbal behavior was not significantly associated with stated impact on receivers. One of the most obvious is that couples may have reacted to the presence of the video camera by toning down and constricting the nonverbal signals accompanying their words. Emotion ordinarily sent by nonverbal means could have been displaced into the verbal channel in reaction to a situation in which couples knew they were being taped. For example, a husband who felt self-conscious in front of the camera might have elected to show disagreement with his wife primarily with words, rather than with his face. In this manner, he would look "well-behaved" for the camera, but still manage to get his point across to his wife. This would be consistent with previous research which has shown that, compared with other nonverbal channels, the face is most adept at concealing emotion under certain circumstances (Buller & Burgoon, 1994; Ekman & Friesen, 1969; Zuckerman & Driver, 1985). In addition, as De Paulo (1992) notes, senders' nonverbal behavior is generally not totally out of their awareness. People can and do attempt to present themselves in certain ways by controlling aspects of their nonverbal presentation. Thus, it is possible that couples in this study were particularly motivated to control and perhaps downplay the intensity of their nonverbal signals.

A second methodological reason why senders' nonverbal behavior may not have correlated with receiver impact may involve the amount of interaction time given couples during the Talk Table procedure. Couples were required to interact for a full minute before the tape recorded message played asking them to rate themselves. Different amounts of interaction time were tested during pilot work and, based on feedback from pilot couples, it was determined that 60 second interaction sequences were the most ecologically valid. (Couples reported that they were unable to carry on meaningful conversations when they were interrupted more than once per minute by the tape recorder).

Since couples based their ratings of each other on one-minute sequences, raters also based their ratings of couples on the same one-minute sequences. Raters were thus required to summarize over 60 seconds worth of nonverbal behavior in order to make their rating decisions. Summarizing over relatively long time intervals may have led raters to overlook some of the subtleties in senders' nonverbal behavior that receivers took into account when rating senders. In particular, it is possible that raters overlooked some of the more subtle, "shorthand" nonverbal cues characteristic of long-term, established relationships (eg., a wink or a slight raise of the eyebrows) that replace the more exaggerated and complex nonverbal displays that predominate in a new relationship (Riggio, 1992). One way to eliminate this problem in future studies would be to continue to have couples rate themselves every minute. Additionally, however, they should ask receivers what aspect of senders' behavior they were responding to most when making the rating (face, voice, or words) and when during the interaction that behavior occurred. The videotape could then be used to determine whether raters agreed with receivers on their assessments or whether receivers were responding to something idiosyncratic in senders' behavior.

Certain methodological aspects of this study, such as the use of the video camera and the length of the interaction time, may have limited the generalizability of the results. However, it is also true that the methodology has a major strength that previous studies on verbal-nonverbal incongruence do not and that may have produced more ecologically valid findings. Specifically, while previous studies on the resolution of incongruent verbal and nonverbal signals have used staged interactions or posed behaviors (e.g., Argyle, Alkema & Gilmour, 1970; Fleming & Darley, 1991; Kahn, 1970; Mehrabian & Wiener, 1967; Noller, 1980; Reilly & Muzekari, 1979; Walker & Trimboli, 1989), this study used spontaneous, nonstaged interactions which more closely approximate naturally occurring dyadic interactions. Therefore, while senders' nonverbal behavior may have been less exaggerated than in previous studies which have directly coached production of nonverbal cues, it also was more true to life. As De Paulo (1992) notes, nonverbal cues which are spontaneously produced, as in the present study, are often less easily interpreted than when deliberately posed. This is consistent with previous research comparing nonverbal expression in naturally occurring and contrived interactions. Such research has found that nonverbal expressions of emotion in natural conversational contexts are more ambiguous and not easily separated from verbal language (Fujita, Harper & Wiens, 1980; Knapp & Hall, 1997; Motley, 1993; Tucker & Riggio, 1988). Similarly, in the case of unrehearsed marital interactions, Noller (1992) has found that couples' nonverbal behavior is less exaggerated and stereotyped when couples engage in free interactions than in interactions with standardized content. What this means is that nonverbal cues in naturally occurring dyadic interactions may not carry as much emotional weight as previously assumed and that verbal language is equally important in conveying emotion. The exception to this might be when one spouse deliberately sends an incongruent message, for example in the case of sarcasm. In this case, because the nonverbal message is deliberately and consciously encoded, it is likely more obvious to the receiving spouse and may thus have more of an emotional impact.

The finding that receivers' ratings correlated more with senders' words than with their nonverbal behavior may have important implications for interpersonal theory, which suggests that emotional messages are conveyed primarily through nonverbal means. Interpersonal theorists have assumed that when verbal and nonverbal signals conflict, nonverbal signals carry the brunt of the emotional message. While this may be the case in laboratory-based studies in which nonverbal signals have been intentionally exaggerated, or in cases in which people are deliberately trying to deceive (e.g., Babad, Bernierni & Rosenthal, 1989; Zuckerman, De Paulo & Rosenthal, 1986), it may not be the case in naturally occurring interactions such as those in the present study. In such interactions, nonverbal signals may be ambiguous or fleeting and may act more as brief emotional "interjections" than entire emotional messages themselves (Motley, 1993; Riggio, 1992). The implications for interpersonal theory and for subsequent research are that nonverbal signals are best studied in terms of their verbal context and that verbal language is important in conveying emotion.

REFERENCES

Argyle, M. Alkema, F. & Gilmore, R. (1971). The communication of friendly and hostile attitudes by verbal and nonverbal signals. *European Journal of Social Psychology, 1,* 385-402

Babad, F., Bernieri, F. & Rosenthal, R. (1989). Nonverbal communication leakage in the behavior of biased and unbiased teachers. *Journal of Personality and Social Psychology, 56,* 89 – 94.

Beier, E., & Young, D. (1984). *The silent language of psychotherapy,* New York: Aldine.

Buller, D.B., & Burgoon, J.K. (1994). Deception: Strategic and nonstrategic behavior. In J. Daly & J. Wiemann (Eds.), *Strategic Interpersonal Communication*, pp. 191-224. Hillsdale, NJ: Erlbaum.

Burgoon, J., Buller, D., & Woodall, W. (1996). *Nonverbal communication: the unspoken dialogue* (2nd ed.). New York: McGraw-Hill.

DePaulo, B. (1992). Nonverbal behavior and self presentation. *Psychological Bulletin, 111,* 203-243.

Capella, T.N. & Palmer, M.T. (1989). The structure and organization of verbal and nonverbal behavior: Data for models of reception. *Journal of Language and Social Psychology, 8,* 167-191.

Duke, M. P., & Nowicki, S. (1982). A social learning theory analysis of interactional theory concepts and a multidimensional model of human interaction constellations. In J. C. Anchin & D. J. Kiesler (Eds.), *Handbook of Interpersonal Psychotherapy* (pp. 78-95). New York: Pergamon Press.

Ekman, P., & Friesen, W.V. (1969). Nonverbal leakage and clues to deception. *Psychiatry, 32,* 88-105.

Filsinger, E., & Wilson, M. (1983). Social anxiety and marital adjustment. *Family Relations Journal of Applied Family and Child Studies, 32,* 513-519.

Fleming, J. & Darley, J. (1991). Mixed messages: the multiple avoidance problem and strategic communication. Special Issue: Social Cognition and Communication: Human Judgment in its Social Context. *Social Cognition, 9,* 25-46.

Fujita, B. N., Harper, R. G., & Wiens, A. N. (1980). Encoding-decoding of nonverbal emotional messages: sex differences in spontaneous and enacted expressions. *Journal of Nonverbal Behavior, 4,* 131-145.

Gottman, J., Notarius, C., Markman, H., Banks, S., Yoppi, B., & Rubin, M. (1976). Behavior exchange theory and marital decision making. *Journal of Personality and Social Psychology, 34,* 14-24.

Gottman, J., Notarius, C., Gonso, J., & Markman, H. (1976). *A couple's guide to communication.* Illinois: Research Press.

Hickson, M., & Stacks, D. (1985). *Nonverbal communication: Studies and applications.* Dubuque, Iowa: W. C. Brown Publishers.

Kahn, M. (1970). Nonverbal communication and marital satisfaction. *Family Process, 9,* 449-456.

Kiesler, D. J. (1979). An interpersonal communication analysis of relationship in psychotherapy. *Psychiatry, 42,* 299.

Kiesler, D. J. (1986). Interpersonal methods of diagnosis and treatment. In J. O. Cavenar (Ed.), *Psychiatry, 1,* 1 – 23. New York: Lippincott.

Kiesler, D. J. (1996). From communications to interpersonal theory: A personal odyssey. *Journal of Personality Assessment, 66,* 267-282.

Knapp, M. & Hall, J. (1997). *Nonverbal communication and human interaction.* New York: Harcourt Brace.

Knudsen, H. R. & Muzekari, L. H. (1983). The effects of verbal statements of context on facial expressions of emotion. *Journal of Nonverbal Behavior, 7,* 202-212.

Mehrabian, A. & Wiener, M. (1967). Decoding of inconsistent communications. *Journal of Personality and Social Psychology, 6,* 108-114.

Mehrabian, A. (1981). *Silent Messages.* Belmont, CA: Wadsworth.

Motley, M. T. (1993). Facial affect and verbal context in conversation. *Human Communication Research, 20,* 3-40.

Noller, P. (1980). Misunderstandings in marital communication: A study of couples' nonverbal communication. *Journal of Personality and Social Psychology, 31,* 1135-1148

Noller, P. (1984). *Nonverbal communication and marital interaction.* International series in experimental psychology, v.9. Elmsford, NY: Pergamon Press.

Noller, P. (1992). Nonverbal communication in marriage. In R. S. Feldman (Ed.), *Applications of Nonverbal Behavioral Theories and Research*. Hillsdale, N. J.: Lawrence Erlbaum.

Noller, P., Feeney, J., Bonnell, D., Callan, V. (1994). A longitudinal study of conflict in early marriage. *Journal of Personal and Social Relationships, 11*, 233-252.

Noller, P., & Gallois, C. (1988). Understanding and misunderstanding in marriage. In P. Noller and M. A. Fitzpatrick (Eds.), *Perspectives on Marital Interaction* (pp. 53 – 77). Clevedon & Philadelphia: Multilingual Matters.

Noller, P., & Guthrie, D. (1989). Assessment and modification of marital communication. Special issue: Behavioral and marital therapy. *Behavior Change, 6*, 124 - 136.

Noller, P. & Venardos, C. (1986). Communication awareness in married couples. *Journal of Social and Personal Relationships, 3*, 31 – 42.

Reilly, S. S. & Muzekari, L. H. (1979). Responses of normal and disturbed adults to mixed messages. *Journal of Abnormal Psychology, 88*, 203-208.

Riggio, R. (1992). Social interaction skills and nonverbal behavior. In R. S. Feldman (Ed.), *Applications of Nonverbal Behavior Theories and Research*, Hillsdale, N. J.: Lawrence Erlbaum.

Satir, V. (1964). *Conjoint Family Therapy*. Palo Alto, California: Science and Behavior Books, Inc.

Satir, V. (1976). Family communication and conjoint marital therapy. In B. N. Ard & C. C. Ard (Eds.), *Handbook of Marriage Counseling*. Palo Alto, CA: Science and Behavior Books.

Schaap C. & Jansen-Nawas, C. (1987). Marital interaction, affect, and conflict resolution. *Sexual and Marital Therapy, 2*, 35-51.

Schachter, J., & O'Leary, K. D. (1985). Affective intent and impact in marital communication. *The American Journal of Family Therapy, 13*, 17- 23.

Sillars, A., Folwell, A., Hill, K. & Maki, B. (1994). Marital communication and the persistence of misunderstanding. *Journal of Personal and Social Relationships, 11*, 611 - 617.

Sullivan, H. S. (1953). *The Interpersonal Theory of Psychiatry*. New York: Norton.

Tucker, J. S. & Riggio, R. E. (1988). The role of social skills in encoding posed and spontaneous facial expressions. *Journal of Nonverbal Behavior, 12*, 87-97.

Van Denburg, T. F., Schmidt, J. A., & Kiesler, D. J. (1992). Interpersonal assessment in counseling and psychotherapy. *Journal of Counseling and Development, 71*, 84-90.

Vangelisti, A. (1994). Couples communication problems: the counselor's perspective. *Journal of Applied Communication Research, 22*, 106-126.

Walker, M., & Trimboli, A. (1989). Communicating affect: the role of verbal and nonverbal content. Third International Conference on Social Psychology and Language. *Journal of Language and Social Psychology, 8*, 229-248.

Zuckerman, M. & Driver, R.E. (1985). Telling lies: Verbal and nonverbal correlates of deception. In A.W. Siegman & S. Feldstein (Eds.), *Multichannel integrations of nonverbal behavior*, pp. 129-148. Hillsdale, NJ: Erlbaum.

Zuckerman, M., Depaulo, B. & Rosenthal, R. (1986). Humans as deceivers and lie detectors. In P. Blanck, R. Buck, and R. Rosenthal (Eds.), *Nonverbal Communication in the Clinical Context.* University Park: Pennsylvania State Press.

The Effects of Confidence and Perception of Test-taking Skills on Performance

Lisa F. Smith
Kean University

This study examined the effects of confidence and perception in test-taking skills on test performance. Participants were 106 undergraduate students in an urban state university, with an average age = 24.42 (*SD* = 6.39) and an average grade point average = 3.22 on a 4.0 scale (*SD* = .44). Each participant completed a demographic questionnaire, the Learning and Study Skills Inventory (Weinstein, Palmer, & Schulte, 1987), a 20 item scale developed for this research to assess attitudes toward test-taking skills (the Self-perceptions of Test Taking Scale), and two multiple choice exams made up of 30 released items from the social studies and mathematics portions of the Praxis Exam (Educational Testing Service, 1992). Regression analyses indicated that confidence was related to test performance but self-perception of test taking skills was not. The results are discussed in terms of assisting students in building confidence in the classroom to bolster test performance.

Students have been known to describe themselves as being "good test takers" or, more commonly, *not* "good test takers." Often, these beliefs are used to excuse poor performances on examinations. Previous research (Pintrich, 1988, 1989; Smith, 2000, 1998a, 1998b; Wolf, 1993; Wolf & Smith, 1995; Wolf, Smith, & Birnbaum, 1995) has shown that when the expectancy-value theory of motivation (Wigfield, 1994) is applied to test performance it can be extended to include personal characteristics, affective components, and emotional reactions. The expectancy-value theory of motivation states that in approaching a task a student takes into account the value of that task and the expectation for success. When applied to test performance, factors such as the perceived amount of effort necessary to successfully complete individual items on the test, personal motivation, and test anxiety must also be considered (Smith, 2000, 1998a, 1998b; Wolf, 1993; Wolf & Smith, 1995; Wolf, Smith, & Birnbaum, 1995). It seems logical that other individual differences contribute to the variability seen in test scores. For example, the level of confidence in the answers one gives on an exam, coupled with an overall positive or negative belief in one's test-taking skills may well play into the performance on that exam.

Author info: Correspondence should be sent to: Dr. Lisa F. Smith, Dept. of Psychology, Kean University, Union, NJ 07083.

North American Journal of Psychology, 2002, Vol. 4, No. 1, 37-50.
© NAJP

In general, the study of affective factors and test performance is fairly new in the literature. The results of current research that examine whether confidence is related to test performance have been varied. Some work has demonstrated that students are fairly accurate in their self-assessments on test-performance. Sharma (1993), for example, found that undergraduate engineering and science majors (*n*=38) were quite accurate when asked to predict how well they thought they would perform on a quiz. Results from this study suggested that the formation of a composite score made up of examinees' confidence ratings and their expected scores can provide a meaningful evaluation index as well as an index of personal psychological satisfaction in the examinees.

In a study with junior high school aged students, Sherman (1980) reported a positive relationship between confidence and test performance on three cognitive tests, the *Scholastic Reading Achievement Math Concepts Test* and the *Vocabulary and Differential Aptitude Tests: Space Relations* tests from the *California Achievement Tests*. In a later study (Pressley & Ghatala, 1989), gender differences in confidence levels were found at the junior high school level. Boys in that study tended to be more confident than girls.

Other studies have indicated that the accuracy of confidence ratings may increase with awareness of ability and the accurate assessment of the difficulty of the material. Shaughnessy (1979) found that all of the undergraduate participants (*n*=49) taking a multiple choice classroom test demonstrated some confidence-judgment accuracy; however, there was a strong positive relationship between confidence judgments and test performance for those participants who had a high level of memory-monitoring ability.

Sjostrom and Marks (1994) reported a similar finding. In their study, 90 undergraduate students in an introductory psychology course rated their confidence in passing each of 12 tests, both before and after each test. Average confidence ratings across all tests were significantly correlated with test performance. When the participants were divided into high, medium, and low performance groups, those in the high group were better able to predict their performance than the other two groups. The authors suggested that this finding indicated that the ratings were probably based more on the participants' assessments of the difficulty of the material on the tests coupled with knowledge of individual ability, than on any perception of individual study skills.

Sinkavich (1995) also found support for Shaughnessy's (1979) findings. This study demonstrated that the level of confidence in the answers given by 67 university students was related to exam scores. Overall, those participants with a high mean confidence rating had significantly higher performance scores on a multiple choice exam than those participants with a low mean confidence rating. Furthermore, the good students were better than the poor students in predicting their test item performance. Sinkavich suggested that this indicated better metamemory accuracy for multiple choice tests for the good students as compared to the poor students.

In fact, overconfidence in performance judgment can lead to problems for poor students. Overconfidence has been shown to be related to test difficulty (Schraw & Roedel, 1994) and belief in quick learning (Schommer, 1990) for this subset of examinees. It seems that these students may both overestimate their test performance ability and believe that they can attain more with less effort than is actually needed to succeed on an examination.

Confidence has also been related to other factors. For example, Savitz (1985) found that the placement of the items on a multiple-choice exam can affect self-confidence evaluations and test performance. When items were randomly ordered on two science exams, junior college students (*n*=427) demonstrated lower levels of confidence and poorer test performance than when easy items were placed at the beginning of the exam. Furthermore, participants made more accurate assessments of their knowledge on the exam that had the easier items at the beginning. In another study, Feldt (1990) found that study time, test strategies, and confidence ratings were significantly related to test performance on factual items, but not on higher order items. Yancey, Humphrey, and Neal (1992) gave 46 undergraduates feedback about their performances on a practice test in a mathematics class. On the subsequent consequential exam, they found a negative relationship between confidence and arousal. Schraw, Dunkle, Bendixen, and Roedel (1995) and Schraw (1997) have argued that confidence in responses to test items is better depicted as domain-general as compared to domain-specific. These two studies have demonstrated that confidence judgments are related not just to performance on a particular test but also to confidence judgments across unrelated tests, including a lexical comparison test, a syllogistic reasoning test, a math test, and a reading test. Based on these findings, Schraw, et al. and Schraw have suggested that metacognitive knowledge about test performance in general affects confidence judgments across domains.

Not all research has found a correlation between confidence and test performance. Glenberg and Epstein (1985) asked participants to rate their confidence in their ability to draw inferences from text, but found no calibration of comprehension. However, a modest increase was observed when the participants were asked to predict future performance. In a later study, Morris (1990) reported that confidence in the ability to answer items on previously studied expository passages was unrelated to test performance.

While there are numerous studies in the literature to indicate that ability, study habits, motivation, and test anxiety are related to performance, there are fewer studies that examine the relationship between personal affective factors and performance. If it can be demonstrated that affective factors such as confidence and belief in test-taking skills are related to test performance, educators can better prepare to work with their students to maximize performance on exams. This study explores the effect of confidence and test-taking skills on performance. The hypotheses for this research are that (a) confidence will be positively related to performance, and (b) self-evaluation of test-taking skills will be positively related to performance.

METHOD

Participants

There were 106 participants in this study. All participants were junior or senior undergraduate students in an urban state university located in central New Jersey. They were enrolled in either introductory statistics (*n* = 51) or experimental psychology (*n* = 55) during the fall 1998 and spring 1999 semesters. The average age of the participants was 24.42, with a standard deviation of 6.39 and a range of 19 to 48. The average grade point average was 3.22 on a 4.0 scale, with a standard deviation of .44 and a range of 2.0 to 4.0. For the completion of all parts of the study, participants were awarded three extra credit points toward their next classroom exam, which equated to one point toward their final average.

Apparatus

The apparatus consisted of a short demographic questionnaire to collect information on variables such as gender, GPA, and age, and the *Learning and Study Skills Inventory* (LASSI) (Weinstein, Palmer, & Schulte, 1987), a 20-item scale to measure attitude toward test-taking skills, and two multiple-choice tests designed for teachers.

The LASSI is made up of 77 items yielding 10 subscales. These subscales evaluate attitude and interest, motivation, time management, anxiety, concentration and attention to academic tasks, information processing, selecting main ideas and recognizing important information, use of support techniques, self-testing and preparation for classes, and test strategies and preparing for tests. The LASSI was selected for the content of the subscales and because it is known to be psychometrically sound. Coefficient alphas for the subscales range from .68 to .86; test-retest coefficients for the subscales range from .72 to .85.

The 20-item scale developed for this research for the purpose of assessing attitudes toward test-taking skills, is called the Self-perceptions of Test Taking Scale (SPOTT) (see Appendix). Two multiple choice exams were given, made up of 30 publicly released items, 15 each taken from the social studies and mathematics portions of the Praxis Exam, (Educational Testing Service, 1992). Items were selected to reflect a wide range of *p*-values as supplied by the test publisher.

Procedure

Participants signed and were given a copy of informed consent. Each participant completed the demographic questionnaire, the LASSI, the test-taking scale, and the multiple choice exam. They were given sufficient but not unlimited time to complete the multiple-choice exam. Most participants took about 30 to 40 minutes to complete the exam. In addition to giving the answer for each item, participants were asked to rate their confidence in their response to each item, on a 1 (low) to 5 (high) scale. Finally, using a 1 (low) to 10 (high) scale, participants were asked to rate their overall test-taking ability. Students self-scored the LASSI and had the opportunity to meet with the instructor to discuss their individual results. After completion of all parts of the study each participant received a debriefing form.

RESULTS

Coefficient alpha was calculated for the SPOTT scale to determine the internal reliability of this instrument. Alpha equaled .69, which indicated that this scale probably contains more than one construct. To examine this, a factor analysis of the SPOTT was completed, using a varimax rotation. A three-factor solution was obtained. The following items from the scale loaded on the first factor:

4. I often have difficulty understanding the instructions on a test.
5. During a test, I am easily distracted and lose my concentration.
6. My anxiety during a test interferes with my ability to focus on the test questions.
7. I often have difficulty figuring out what the test questions are asking.
8. I often get hung up on a question that I don't know, rather than going on to the next question.
12. When taking a test, I often "go blank" on material I studied and thought I knew well.
13. I often feel rushed while taking a test, like I will run out of time.

This factor appears to concern test anxiety. The items that loaded on the second factor were (an asterisk indicates that the item was reversed for scoring):

*1. When taking a test, I feel that I am in control.
*2. I try to get myself psyched up (motivated) just before a test.
*3. I usually begin a test feeling confident that I can do well.
7. I often realize that I have studied the wrong material.
13. I get very tired while taking tests.
*18. When I finish a test, I usually have a good idea of how well I did.
*20. I think I have good test-taking strategies/skills.

This factor seems to relate to confidence in test taking. The items that loaded on the third factor were (an asterisk indicates that the item was reversed for scoring):

*10. I look through the whole test before I begin.
*11. I leave the most difficult questions on a test for last.
15. On multiple choice questions, I often get down to two of the options and then can't decide between them.
16. On short answer or essay questions, I write everything I can think of on the topic instead of focusing on what is being asked by the question.
17. I often think the test questions do not reflect the material that was emphasized in class.
18. When I take a guess on a question, I always seem to make the wrong choice.

This factor appears to concern test-taking behaviors. As the three factors all pertain to perceptional aspects of test-taking that may affect performance, the total scale score was used for all subsequent analyses.

Next, cross-tabulations were computed for each item with that item's confidence score. The items were scored dichotomously; the confidence ratings were on a 1 (no confidence) to 5 (absolutely confident) scale. Two pieces of information were extracted from these crosstabulations. The first examined the percent of those participants who were absolutely confident in their response but got the item wrong, as well as the percent of those participants who indicated no confidence in their response but got the item correct. In other words, the perception of confidence was not congruent with performance on that item. Next, each cross-tab was examined to compare the p value for this sample to the empirical *p* value of the item. In this way, potential "trick" items could be identified. See Table 1 for these results. There were three items that were identified as potential trick items for this sample. The first was a social studies item that asked the examinee to recognize a situation that would identify the term "détente." The second was a social studies item that required selecting the main idea of a short descriptive paragraph. The last was a mathematics item that required knowledge of fractions.

Table 1 Comparison of Item Responses to Confidence Ratings and Empirical *p* Values

Item	Percent Confident/Wrong	Percent Not Confident/Correct	Observed p	Empirical p
Social Studies 1	.9	.9	.85	.73
Social Studies 2	5.8	4.8	.75	.44
Social Studies 3	1.0	21.2	.28	.53 *
Social Studies 4	4.8	1.9	.62	.66
Social Studies 5	7.5	1.9	.06	.15
Social Studies 6	3.9	2.9	.73	.85
Social Studies 7	2.9	5.8	.56	.56
Social Studies 8	3.8	4.8	.35	.46
Social Studies 9	---	4.7	.63	.81 *
Social Studies 10	1.9	6.6	.49	.61
Social Studies 11	1.9	3.8	.79	.77
Social Studies 12	1.9	14.2	.65	.64
Social Studies 13	1.0	5.8	.62	.65
Social Studies 14	7.8	8.7	.42	.47
Social Studies 15	1.0	3.9	.71	.78
Math 1	14.6	---	.70	.73
Math 2	4.9	2.9	.85	.71
Math 3	10.4	5.7	.83	.83
Math 4	4.8	2.9	.72	.79
Math 5	11.4	10.5	.59	.67
Math 6	9.5	4.8	.83	.89
Math 7	17.3	2.9	.16	.11
Math 8	6.7	3.8	.73	.78
Math 9	6.7	10.5	.49	.59
Math 10	4.8	7.7	.48	.63 *
Math 11	1.0	8.7	.40	.53
Math 12	6.8	6.8	.37	.47
Math 13	3.0	3.0	.55	.62
Math 14	2.9	14.4	.35	.38
Math 15	34.3	---	.14	.16

Note. * indicates potential trick question for this sample

The next analyses used two regressions to determine which of the affective variables in the study were related to performance. The dependent variable in the first regression was the Total Correct score and the dependent variable in the second regression was the Total Confidence score. Analyses using all affective items in the study, including the LASSI scores, did not yield significant results for any of the LASSI subscales for the regression using Total Correct as the dependent variable. Therefore, the regression for Total Correct was re-run using grade point average, confidence score, SPOTT score, and self-rating (1 = low, 10 = high) of test-taking skills as the predictors. The *R* was .499 and the *R* square was .249. Results for the model were $F_{(4)} = 7.362, p < .001$. Confidence and grade point average were significant predictors. The regression coefficients are shown in Table 2.

TABLE 2 Results of Regression for Total Correct

Predictor	Beta	*t*	*p*
Self-rating of Test Taking Skills	-.043	-.393	.696
Confidence Score	.431	4.186	.001
SPOTT Score	-.030	-.282	.778
Grade Point Average	.207	2.017	.047

An initial regression using the Total Confidence score as the dependent variable indicated that the LASSI subscale for motivation should be included in the model. Therefore, the predictors for this regression were self-rating of test taking skills, SPOTT score, the LASSI motivation subscale score, and grade point average. The *R* was .500 and the *R* square was .25. The model was significant, $F_{(4)} = 7.40, p < .001$. Self-rating of test taking skills, the LASSI motivation subscale score, and grade point average were significant predictors. The regression coefficients are shown in Table 3.

TABLE 3 Results of Regression for Confidence

Predictor	Beta	*t*	*p*
Self-rating of Test Taking Skills	.342	3.33	.001
LASSI Motivation Subscale Score	-.244	-2.38	.019
SPOTT Score	.116	1.08	.281
Grade Point Average	.215	2.01	.048

Although the SPOTT did not prove to be a significant predictor in either regression, it was correlated with grade point average as well as several of the LASSI subscales and the self-rating and confidence scores (see Table 4).

Finally, correlations were computed to directly address the hypotheses that confidence would be positively related to performance and self-evaluation of test taking skills would be positively related to performance. There was a significant correlation for confidence and total score, $r = .46$, $p < .001$, lending support for the first hypothesis. The measures of test taking skills, however, did not yield significant correlations with performance. The correlation of the SPOTT with total score was $r = .135$, $p = .167$ and the correlation for the self-rating of test taking and total score was $r = .139$, $p = .159$. It was interesting to note that, for this sample, the correlation between the LASSI subscale for test taking strategies and total score was not significant, $r = .149$, $p = .127$.

TABLE 4 Correlation of SPOTT Score with Selected Variables

Variable	*r*
Grade Point Average	.26 **
LASSI Attitude and Interest	.30 **
LASSI Motivation	.30 **
LASSI Time Management	.25 **
LASSI Anxiety	.69 ***
LASSI Concentration	.45 ***
LASSI Select Main Ideas	.48 ***
LASSI Test Strategies	.67 ***
Self-rating	.41 ***
Confidence Score	.25 **

Note. *** $p < .001$ ** $p < .01$

DISCUSSION

This study set out to examine the effects of confidence and self-perception of test taking skills on performance. The results indicated that although confidence was significantly related to performance for this sample, self-perception of test taking skills was not significantly related to test performance. Prior research findings with regard to confidence in responses to test items have been mixed. A full range of results have been reported, from a significant relationship between the level of confidence in the answers given and exam scores (Sinkavich, 1995) to no correlation between confidence and test performance (Morris, 1990), with a curvilinear relationship (Yancey, Humphrey, & Neal, 1992) falling between the extremes. No studies were found that examined confidence within the larger framework of personal beliefs about test taking skills.

There is sufficient evidence from this study to suggest that working in the classroom to increase confidence may have a positive effect on test performance. Classroom tasks that assist students in their preparation for examinations or that familiarize students with test formats under less stressful or less consequential settings than exams may bolster confidence for actual examinations. Research (Kellaghan, Madaus, &

Raczek, 1996; MacIver, Stipek, & Daniels, 1991) has indicated that increased confidence leads to increased effort and valuing of education in general. On the other hand, it is interesting to note that for almost every item on the test there were at least some students who were highly confident that they had correctly answered the item, when in fact they had not. (The opposite of this was also true; that is, completely unconfident students got items correct, but that may be the result of fortunate guessing.) This finding might be construed as overconfidence from the perspective of the student or as a trick question on the test (fooling students into believing they had the correct answer, when in fact, they fell for the "trick") or even as reading too much into an item. In any case, the role of confidence in test performance is an intriguing one. Confidence in one's abilities, or lack thereof, affects how an individual approaches each new item on the test. The relationship between confidence and performance can be cumulative as one progresses through the examination. This is a particularly interesting possibility for continued research along this line.

In terms of the interpretation of test scores, it seems that consideration of affective factors such as confidence and motivation would allow a more complete and accurate evaluation of student performance. The results of this study indicate that confidence is a strong predictor of performance and in turn is predicted by affective measures such as self-perception of test-taking skills and cognitive measures such as grade point average. It is particularly interesting to note that the single item measure of self-rating of test-taking ability (asked prior to taking the test) correlated at .462 with the summation of the confidence scores given on an item-by-item basis, but only .139 with the total score (number right). Confidence while taking a test appears to be a function of both how well the student perceives he or she is doing during the test and a general tendency to be confident in such situations. To continue this line of argument, one can think of confidence as a positive analogue or "flip-side" to anxiety, and to consider the item-by-item fluctuations in confidence as a "state-confidence" which would be equivalent to "state-anxiety" and the more enduring estimate of confidence in performance given before the test as a "trait-confidence." Of course, this is just speculative at this time, but it certainly seems to be the case that confidence has a role in test performance, perhaps a very useful one.

The demographics of the sample, the subject matter of the exams, and the fact that the exam itself was not part of the course grade limit the generalizability of the findings. This may not be as much of a problem given Schraw et al. (1995) and Schraw's (1997) findings that confidence judgments generalize across domains or if additional research supports the notion of "trait-confidence." Further examination of this area is needed to help determine the components of a test score. Most assuredly, a consideration of the affect for the person taking an exam must be factored into any complete understanding of a test score.

REFERENCES

Educational Testing Service. (1992). *NTE® Core Battery Tests Practice & Review* (pp. 26-43). Princeton, NJ: Educational Testing Service.

Feldt, R. C. (1990). Test expectancy and performance on factual and higher-level questions. *Contemporary Educational Psychology, 15*, 212-223.

Glenberg, A. M. & Epstein, W. (1985). Calibration of comprehension. *Journal of Experimental Psychology Learning Memory and Cognition, 11*, 702-718.

Kellaghan, T., Madaus, G. F., & Raczek, A. (1996). *The use of external examinations to improve student motivation*. Washington, DC: The American Educational Research Association.

MacIver, D. J., Stipek, D. J., & Daniels, D. H. (1991). Explaining within-semester changes in student effort in junior high school and senior high school courses. *Journal of Educational Psychology, 83*, 201-211.

Morris, C. C. (1990). Retrieval processes underlying confidence in comprehension judgments. *Journal of Experimental Psychology - Learning, Memory, and Cognition, 16*, 223-232.

Pintrich, P.R. (1988). A process-oriented view of student motivation and cognition. In J.S. Stark & R. Mets (Eds.), *Improving teaching and learning through research* (pp. 55-70). San Francisco: Jossey-Bass.

Pintrich, P.R. (1989) The dynamic interplay of student motivation and cognition in the college classroom. In C. Ames & M. Maehr (Eds.), *Advances in achievement and motivation* (Vol. 6, pp. 117-160). Greenwich, CT: JAI Press.

Pressley, M., & Ghatala, E. S. (1989). Metacognitive benefits of taking a test for children and young adolescents. *Journal of Experimental Child Psychology, 47*, 430-450.

Savitz, F. R. (1985). Effects of easy examination questions placed at the beginning of science multiple-choice examinations. *Journal of Instructional Psychology, 12*, 6-10.

Schommer, M. (1990). Effects of beliefs about the nature of knowledge on comprehension. *Journal of Educational Psychology, 82*, 498-504.

Schraw, G. (1997). The effect of generalized metacognitive knowledge on test performance and confidence judgments. *Journal of Experimental Education, 65*, 135-146.

Schraw, G., Dunkle, M. E., Bendixen, L. D., & Roedel, T. D. (1995). Does a general monitoring skill exist? *Journal of Educational Psychology, 87*, 433-444.

Schraw, G. & Roedel, T. D. (1994). Test difficulty and judgment bias. *Memory and Cognition, 22*, 63-69.

Sharma, N. K. (1993). How confident are humans about their self-assessment? Special issue: Psychoeducational assessment. *Psychological Studies, 38*, 150-156.

Shaughnessy, J. J. (1979). Confidence-judgment accuracy as a predictor of test performance. *Journal of Research in Personality, 13*, 505-514.

Sherman, J. A. (1980). Predicting mathematics grades of high school girls and boys: A further study. *Contemporary Educational Psychology, 5*, 249-255.

Sinkavich, F. J. (1995). Performance and metamemory: Do students know what they don't know? *Journal of Instructional Psychology, 22*, 77-87.

Sjostrom, K. P. & Marks, A. (1994). Pretest and posttest confidence ratings in test performance by low-, medium-, and high-scoring students. *Teaching of Psychology, 21*, 12-16.

Smith, L. F. *Confidence and test performance*. Paper presentation at the annual meeting of the American Educational Research Association, New Orleans, April, 2000.

Smith, L.F. (1998a). *The effects of difficulty, consequence, and effort on performance*. Paper submitted in partial fulfillment of Kean University Untenured Faculty/Librarian Research Initiative.

Smith, L.F. (1998b). *The influence of test consequence and perceived item effort*. Paper submitted in partial fulfillment of Kean University Untenured Faculty/Librarian Research Initiative.

Wigfield, A. (1994). Expectancy-value theory of achievement motivation: A developmental perspective. *Educational Psychology Review, 12*, 265-310.

Wolf, L.F. (1993). The effects of motivation and anxiety on test performance. *Dissertation Abstracts International, 54*, 04A. (University Microfilms No. 93-24654)

Wolf, L.F., & Smith, J.K. (1995). The consequence of consequence: Motivation, anxiety and test performance. *Applied Measurement in Education, 8*, 227-242.

Wolf, L.F., Smith, J.K. & Birnbaum, M.E. (1995). Consequence of performance, test motivation, and mentally taxing items. *Applied Measurement, 8*, 341-351.

Yancey, G. B., Humphrey, E., & Neal, K. (1992). How perceived incentive, task confidence, and arousal influence performance. *Perceptual and Motor Skills, 74*, 279-285.

APPENDIX
Self-Perceptions of Test Taking

Please respond to each of the following statements on a scale of 1 (strongly disagree) to 5 (strongly agree).

1. When taking a test, I feel that I am in control.
2. I try to get myself psyched up (motivated) just before a test.
3. I usually begin a test feeling confident that I can do well.
4. I often have difficulty understanding the instructions on a test.
5. During a test, I am easily distracted and lose my concentration.
6. My anxiety during a test interferes with my ability to focus on the test questions.
7. I often realize that I have studied the wrong material.
8. I often have difficulty figuring out what the test questions are asking.
9. I often get hung up on a question that I don't know, rather than going on to the next question.
10. I look through the whole test before I begin.
11. I leave the most difficult questions on a test for last.
12. When taking a test, I often "go blank" on material I studied and thought I knew well.
13. I get very tired while taking tests.
14. I often feel rushed while taking a test, like I will run out of time.
15. On multiple choice questions, I often get down to two of the options and then can't decide between them.
16. On short answer or essay questions, I write everything I can think of on the topic instead of focusing on what is being asked by the question.
17. I often think the test questions do not reflect the material that was emphasized in class.
18. When I finish a test, I usually have a good idea of how well I did.
19. When I take a guess on a question, I always seem to make the wrong choice.
20. I think I have good test-taking strategies/skills.

Stimulus Load and Age Effects in Face Recognition: A Comparison of Children and Adults

Mitchell M. Metzger
The Pennsylvania State University, Shenango Campus

This experiment assessed the effects of stimulus load and age in face recognition. As both factors have been implicated as important determinants in face recognition (Shapiro & Penrod, 1986), this experiment sought to examine the potential interactive role of these variables. That is, the design examined whether increased stimulus load (i.e., the number of stimuli viewed during study and recognition testing) would impair the performance of young children more than adults. Participants were divided into three age groups (children, college students, and middle-aged adults) and were given a face recognition test. Subjects were further divided by level of test difficulty, as stimulus load was manipulated across three conditions. Results indicate that the increased stimulus load caused lower hit rate and d-prime scores and increased false alarm scores in all subjects. Results also showed that both groups of adults had higher hit rate and d-prime scores compared to children. However, the lack of a statistically significant interaction between these variables suggests that the increase stimulus load affected all subjects equally. Further analyses suggest that the differences among groups were not attributable to changes in response criterion, as all groups demonstrated similar levels of response bias. Implications for age influences in face recognition are discussed.

The study of face identification has grown over the past several decades, and a voluminous body of literature has been accumulated. In an attempt to identify key factors that influence face recognition, Shapiro and Penrod (1986) conducted a meta-analysis that examined over 100 studies containing almost 1000 experimental conditions. From this analysis, a number of factors emerged as variables important in face memory. Among many other variables, the Shapiro and Penrod (1986) meta-analysis confirmed that the participants' age and stimulus load emerged as factors that influence face recognition. In their analysis, stimulus load was defined as a combination of exposure time and the number of faces present during familiarization and recognition testing. In essence, their analysis confirmed that (1) performance on face

---

Author info: Correspondence should be sent to: Mitchell Metzger, Ph.D.The Pennsylvania State University, Shenango campus 147 Shenango Ave. Sharon, PA 16146 Email: mxm69@psu.edu

North American Journal of Psychology, 2002, Vol. 4, No. 1, 51-62.

© NAJP

recognition tasks improves with age, and (2) performance decreases as the difficulty of the task increases (e.g., as the number of faces to be recognized increases). These findings come as little surprise, considering the seemingly robust effect of these two variables in face recognition (Carey, 1992; Flin, 1980; Podd, 1990; Shapiro & Penrod, 1986). However, to date, the possible interactive effect of these variables has yet to be examined. Is it possible that the performance of young children is affected more than that of adults when the difficulty of the face recognition task is manipulated by altering stimulus load? There is ample evidence demonstrating that certain variables in face recognition do affect children and adults differently. In a recent face recognition study, Johnston and Ellis (1995) reported that adults performed much better than children when identifying "distinctive", rather than "typical", faces. However, they also reported that, contrary to adults, young children did not show superior performance for distinctive faces.

There also appears to be little consistency across experiments in the number of faces viewed during face recognition tests, and this makes cross-experiment comparisons difficult. For example, Sophian and Stigler (1981) tested young children and adults on a face recognition task in which participants were required to identify 12 faces that had been slightly altered by the manipulation of eye, mouth, and nose features. Johnston and Ellis (1995) tested children and adults on a face recognition task, and these participants were required to identify 18 target faces from a random array of 36 target and foil (distractor) faces. This inconsistency of target and distractor ratios is evident in many face recognition studies, as mentioned by Podd (1990) who stated that "considerable variation does occur" (p. 47). As the majority of face recognition studies show clear adult superiority over children on performance, it is often difficult to make comparisons across studies because of such variations in methodology. For instance, perhaps it is more difficult to identify 18 faces from an array of 36 faces than to identify 12 faces that have been slightly changed. The Shapiro and Penrod (1986) meta-analysis confirmed that stimulus load is an important determinant in face recognition, but this factor should be explored more fully in order to gain a complete understanding of the differences between children and adults on face recognition. Aside from a single study that tested young adults (Podd, 1990), little attention has been given to the study of stimulus load effects in face recognition. In addition, no studies have been conducted on the effects of stimulus load for subjects of different ages.

To accomplish this aim, an experiment was conducted to determine the effects of increasing stimulus load on the performance of children compared to young adult (aged 18-33) and middle-aged adult (aged 40-67) subjects. In essence, the three groups of participants completed a face recognition task that varied the level of difficulty by manipulating stimulus load (number of faces viewed during study and recognition testing). As Johnston and Ellis (1995) observed differential age effects for distinctiveness in face recognition, the present experiment sought to determine whether stimulus load would also differentially affect children and adults. If there is a differential age effect of stimulus load in this procedure, a significant interaction between the age and stimulus load variables should be evident. That is, if an increased stimulus load impairs children's recognition more than adults, their performance should decline more rapidly than the performance of adults as the task difficulty increases. However, if increasing stimulus load affects both children and adults similarly, decreases in performance should be observed in all groups with no interaction between stimulus load and age.

METHOD

Participants

Participants in this experiment were divided into three groups, based on age. Thirty-six children (20 male and 16 female) between the ages of 8-10 (*M* = 9.3) were recruited from a suburban elementary school and were assigned to the CHILD group. Thirty-six college students (17 male and 19 female) between the ages of 18-33 (*M* = 22.0) were assigned to the COLLEGE group, and 36 participants (17 male and 19 female) between the ages of 40-67 (*M* = 49.6) were assigned to the ADULT group. The ADULT group consisted of students, faculty, and staff at Pennsylvania State University, as well as individuals recruited from church groups and local organizations. College students received extra course credit in Introductory Psychology for participation, while the participants in the CHILD and ADULT groups received no compensation.

Materials

Face stimuli consisted of composites created with the ComPhotoFit facial construction software (Sirchie, Inc., Youngsville, NC), described earlier by Metzger (1999). With this software, the completed composite appears to be an actual black and white photograph of a person and not an artistic rendering or police sketch. The stimuli created for this experiment consisted of young Caucasian males, all with short, dark hair. None of the faces contained eyeglasses or other unique, identifying features. The composite images were incorporated into a Microsoft PowerPoint slideshow. The participants in the CHILD group were tested individually and viewed the stimuli on a 40-cm computer monitor. The size of each image was approximately 7.5 x 12.5-cm, and all of the subjects in the CHILD group viewed the images at a distance of 30 to 60-

cm from directly in front of the computer screen. Participants in the COLLEGE and ADULT groups viewed the facial stimuli on a projector screen in a room with elevated seating. The images were cast on the screen with a LCD projector, and the faces were approximately 60 × 120-cm. Subjects in these conditions were tested in groups of 2-12, and were seated approximately 4.5-7.5-m from the screen. Subjects in all three conditions had a clear view of the stimuli, regardless of whether tested individually or as a group. Ideally, the children would have been tested in a group setting similar to the adult participants. However, because of a lack of appropriate facilities at the elementary school, an individual testing procedure for the children was adopted. Figure 1 illustrates a sample composite used in this experiment.

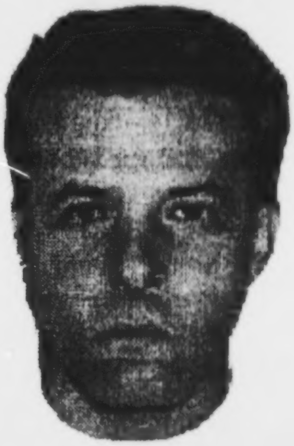

FIGURE 1 A sample composite face used in this experiment. This computer-generated composite was created by blending separate hair, eyes, nose, mouth, and chin features.

Sixty faces were selected from a database of composite pictures and then randomly divided into two sets of 30, while one set served as targets and the other set served as distracters. With these two sets of faces three different slideshows were constructed for this experiment, with each presentation containing a different number of target stimuli to be viewed during the study phase (10, 20, or 30) and recognition testing (20, 40, or 60). For subjects viewing 10 faces during study (and 20 during testing), faces 1-10 from the target group were the targets, and faces 1-10 from the distractor group were the distractors. For the recognition test, these 20 faces were randomly mixed together and presented to subjects. For subjects viewing 20 faces at study (and 40 faces during testing), faces 11-20 from the target group were added to faces 1-10 that were viewed by subjects seeing 10 faces during study. To complete the sequence of 40 faces necessary for the test, target faces 11-20 were randomly mixed with distractor faces 11-20 and added to the end of the sequence of faces viewed by subjects in the previous group. In short, the stimuli viewed by subjects seeing 20 faces at recognition served as the first 20 faces for those subjects viewing 40 faces at recognition. To complete the final slideshow (for subjects seeing 30 faces at study and 60 at test), the same procedure was adopted, as target and distractor faces 21-30 were added to the previous slideshow. This method for constructing the slideshows was taken from an earlier experiment assessing stimulus/memory load in face recognition, which allowed the initial twenty photographs viewed during recognition to be identical, regardless of group assignment (Podd, 1990).

Procedure

The methodology for this experiment consisted of study and recognition phases. During the study phase, subjects viewed each of the faces (10, 20, or 30) for 5-s, with a 2-s inter-stimulus interval (ISI), during which time the screen was blank. Response sheets were then distributed and subjects were given instructions for the recognition test. During recognition, subjects viewed a random presentation of the previously-seen targets and an equal number of distractor stimuli that were unfamiliar (20, 40, or 60 faces total). During the recognition phase, target probability was .5. Each face appeared on the screen for 5-s with a 2-s ISI (blank screen), and subjects responded to whether each face was "old" (a previously-seen target) or "new" (a face they had not seen before). Subjects were instructed to guess if they were not sure of the correct response.

Three dependent variables were measured: Hit Rate (HR-proportion of correctly identified targets), False Alarm Rate (FAR-proportion of distractors identified as targets) and effect size (*d'*). In addition to these recognition indexes, response bias (*c*) was also measured, which assesses subjects' willingness to report a face as "old" or "new" (MacMillan & Creelman, 1991). *C* (which ranges between -2.33 and +2.33) scores near zero indicate no bias, positive scores indicate a bias to report faces as "new", and negative scores indicate a bias to report faces as "old". Measurement of *c* allows one to determine response bias changes even though sensitivity (*d'*) may remain constant. In other words, do subjects perform poorly because they have a bias to respond "old" or "new", or do they perform poorly because their memory for faces has been affected? Should children perform more poorly than adults, for instance, this statistic would indicate whether the performance decrement is the result of a difference in memory, or a difference in the willingness to report faces as "old" or "new".

Design

Three levels of stimulus load during study/recognition were varied in this experiment (10/20, 20/40, and 30/60), and three age levels were tested (CHILD, COLLEGE, and ADULT groups). This produced a 3 × 3 factorial design with both stimulus load and age as between group variables. Each group contained an equal number of subjects (n=12), and there were approximately equal proportions of males and females assigned to each condition.

RESULTS

The mean values for HR, FAR, d', and c for all nine experimental conditions are presented in Table 1. Separate three-way ANOVAs were computed on these values, and pair-wise comparisons were analyzed with Fisher's LSD tests.

TABLE 1 Mean HR, FAR, d', and c Scores of the CHILD, COLLEGE, and ADULT Groups in the 10/20, 20/40, and 30/60 Conditions. (SDs in parentheses).

GROUP	10/20	20/40	30/60
Hit Rate			
CHILD	.71 (.16)	.62 (.11)	.62 (.08)
COLLEGE	.79 (.14)	.74 (.07)	.72 (.12)
ADULT	.86 (.09)	.79 (.15)	.68 (.13)
False Alarm Rate			
CHILD	.16 (.07)	.23 (.11)	.25 (.11)
COLLEGE	.17 (.12)	.19 (.10)	.26 (.10)
ADULT	.18 (.18)	.18 (.10)	.24 (.16)
d'			
CHILD	1.68 (.73)	1.11 (.63)	1.09 (.58)
COLLEGE	2.02 (.80)	1.62 (.62)	1.29 (.36)
ADULT	2.45 (1.16)	1.89 (.90)	1.30 (.81)
c (*response bias*)			
CHILD	.19 (.32)	.25 (.18)	.22 (.31)
COLLEGE	.07 (.42)	.16 (.19)	.03 (.31)
ADULT	.02 (.48)	.06 (.22)	.16 (.28)

Statistical analysis of HR yielded significant main effects for both stimulus load [$F_{(2.99)} = 8.04$, $p<.001$] and age [$F_{(2.99)} = 10.97$, $p<.0001$], but no significant interaction between the two [$F_{(4,99)}$ = .99, p=.42]. Further analysis confirmed that subjects in the 10-target condition showed higher HR than subjects in either the 20-target or 30-target conditions (p<.05), while subjects in the 20-target and 30-target conditions performed at a comparable level (p>.05). Additionally, both the COLLEGE and ADULT groups had higher HR scores than the CHILD group (p<.05), while there were no reliable differences between the COLLEGE and ADULT group for hit rate (p>.05).

A comparison of the FAR measure produced a different pattern of results, as stimulus load was the only variable that produced a significant effect on FAR [$F_{(2,99)}$ = 4.12, p=.02]. The effects of age [$F_{(2,99)}$ = .12, p=.88] and the stimulus load × age interaction [$F_{(4,99)}$ = .27, p=.90] failed to reach a level of statistical reliability. The significant effect of stimulus load was based on subjects in the 30-target condition exhibiting significantly more false alarms than subjects in the 10-target condition (p<.05). There were no reliable differences between subjects in the 10-target and 20-target conditions or between subjects in the 20-target and 30-target conditions (p>.05).

The pattern of results obtained on d' scores is similar to those calculated for HR. Significant main effects were observed on the stimulus load [$F_{(2,99)}$ = 10.75, p=.0001] and age variables [$F_{(2,99)}$ = 5.37, p=.006]; however, the load × age interaction was not significant [$F_{(4,99)}$ = .62, p=.65]. Subjects in the 10-target condition demonstrated higher d' scores than subjects in either the 20- or 30-target conditions (p<.05), while subjects in the 20- and 30-target conditions demonstrated an equal level of performance (p>.05). For the age variable the CHILD group showed lower d' scores than either the COLLEGE or ADULT groups (p<.05), while the two adult groups exhibited a similar level of performance (p>.05).

An analysis of response bias (c) was used as a manipulation check to determine if it accounted for any differences observed among the groups. That is, were the differences between groups attributable to differences in memory or simply differences in a participant's willingness to report a face as "old" or "new"? The analysis of c indicated that there were no significant differences among any of the groups for age [$F_{(2,99)}$ = 2.13, p = .13], stimulus load [$F_{(2,99)}$ = .43, p = .65], or the age × stimulus load interaction [$F_{(2,99)}$ = .41, p = .81]. While the children (M = .22) demonstrated numerically higher c scores, this bias was not reliably larger than the response bias of the COLLEGE (M = .09) or ADULT (M = .08) group. Additionally, response bias seemed unaffected by stimulus load in this experiment as subjects in the 10 target (M = .09), 20 target (M = .15), and 30 target (M = .14) conditions exhibited similar c scores as stimulus load increased. Similar to the results for the age variable, these results suggest that the decreased performance on HR, FAR, and *d′* as stimulus load increased cannot be attributed to changes in response criterion, but can rather be interpreted as being influenced by changes in memory processes.

DISCUSSION

This experiment varied three levels of stimulus load across child, young adult, and middle-aged adult participants. For HR and *d′* scores, significant main effects were observed for stimulus load and age, but a lack of an interaction suggests that all three age groups were affected equally by the increase in stimulus load. Additionally, the only factor that produced a significant effect on FAR scores was stimulus load, while age and the interaction between the two were not significant. Analyses of response bias (*c*) suggest that these main effects can be interpreted as changes in memory function rather than shifts in decision criterion between the groups. Taken as a whole, this pattern of results suggests two trends. First, the data from the current experiment replicate previous findings (Podd, 1990), as increasing the stimulus load produced a reliable performance decrement in all groups. This performance decrement was independent of age, as all three age groups in the present study demonstrated lower HR and *d′* scores and higher FAR scores as stimulus load increased. Secondly, the CHILD group performed more poorly than adult participants on the facial recognition task, as reflected in lower HR and *d′* scores. This second finding is well founded, as numerous studies have demonstrated that young children typically perform more poorly than adults on tasks of this nature (Carey, 1992; Flin, 1980). However, the impetus for this experiment was to determine if an increased stimulus load would differentially affect subjects of different ages. That is, would an increase in stimulus load have a larger impact on children than adults? The lack of any significant interactions between stimulus load and age suggest there is no differential age effect.

These data suggest that young children are less likely than adults to identify faces they have previously seen, as reflected in lower HR scores for children. However, they also suggest that nine year-old children were no more likely than adult subjects to falsely identify a new face as one they have seen before. Importantly, this second finding was independent of task difficulty (i.e., increased stimulus load). As Davies (1996) has observed, there is little available data on the performance of children in situations where large numbers of facial stimuli are examined. At least for the highest level of stimulus load varied in the present study (30 faces at familiarization/60 faces during recognition), the children and both groups of adults demonstrated similar FAR scores. Further, since children's response bias was similar to those of adults, they were no more likely to employ a biased criterion in identifying faces. The findings from the response bias analyses are encouraging, suggesting that children employed an adult strategy in performing face recognition. That is, when compared to adults, 9-year olds did not employ an overly conservative or liberal approach in the identification of previously-seen faces.

It is interesting to note that significant effects of age occurred on both the HR and d' measures, but not for FAR. This finding appears to be inconsistent with previous reports of age differences for the false alarm measure (Johnston & Ellis, 1995). However, there are several important differences between the two studies that may account for this difference in results. Johnston and Ellis (1995) examined several different factors in their study that were not manipulated in the present experiment, such as distinctiveness and viewing time. They also studied a wider age-range of children. In their study, children of 5, 7, 9, 11, and 13 years were compared to college students. While significant main effects of age were reported for false alarms, it appears that this effect was much more robust in the younger sample of children (5 and 7 year olds). A comparison of the 9-year old children (the average age of subjects in the present experiment) suggests that this group demonstrated a number of false alarms that was comparable to the adult group. The FAR data from Johnston and Ellis (1995) for the 9-year old group does appear to parallel results found in the present experiment.

An alternative interpretation of these data is that a significant age × stimulus load interaction might have emerged if more stimuli were used during the study and recognition phases. That is, it is possible that differential age effects may have been observed if subjects would have viewed more faces during study and recognition. While this is a possibility, the reason for only showing subjects a maximum of 30 faces during study (and 60 during recognition) was due to a concern of attention/fatigue effects in the children. That is, with large stimulus sets, it becomes difficult to sort out the difference between a performance decrement caused by taxing a participant's memory and their becoming bored or fatigued from the procedure. Adults may be capable of holding their attention for periods of 10-15 minutes (Podd's subjects studied 50 faces and identified them from 100). However, the present concern was that the performance of children would have been affected by attention and fatigue effects with increasing stimulus sets. Enns (1990) has reported that younger children do not focus their attention as effectively as adults, so the stimulus set in the present study may not have been large enough to produce an effect. Granted, a larger stimulus set might have produced a significant age × stimulus load interaction; however, within the conditions varied in this experiment, these data indicate the absence of a differential age effect. Future research should examine greater stimulus loads.

The present experiment demonstrated that increased stimulus load affected children, young adults, and middle-aged adults equally during a face recognition task. Additionally, children's hit rates were poorer than both groups of adults, and these group differences cannot be interpreted as changes in response criterion. Although no significant interaction was discovered among the variables studied, this experiment provides data from three diverse age groups, and provides additional insight on the ability of young children to recognize faces from relatively large stimulus sets. While children showed lower HR and *d′* than both groups of adults in the present study, this study confirmed that the young subjects were not differentially impaired compared to adults when the difficulty of the task was increased. This finding suggests several important implications. As known from previous research (Flin, 1980; Johnston & Ellis, 1995), at 9-years of age the face recognition system is not yet fully developed, as there were observed differences between the groups for HR and *d′*. However, 9-year old children and both groups of adults were affected equally by increasing task difficulty (at least the levels studied here), suggesting that some aspects of the face recognition system in children are comparable to the adult subject. Further investigations into the variables of stimulus load and age will assist in advancing our understanding of these important determinants of face recognition.

REFERENCES

Carey, S. (1992). Becoming a face expert. *Philosophical Transactions of the Royal Society of London, Series B, 335*, 95-103.

Davies, G.M. (1996). Children's identification evidence. In S.L. Sporer, R.S. Malpass, & G. Koehnken (Eds.) *Psychological Issues in Eyewitness Identification*. Mahwah: Lawrence Erlbaum Associates.

Enns, J.T. (1990). Relations between components of visual attention. In J.T. Enns (Ed.), *The development of attention: Research and theory*. Amsterdam: Elsevier.

Flin, R. (1980). Age effects in children's memory for unfamiliar faces. *Developmental Psychology, 16*(4), 373-374.

Johnston, R.A., & Ellis, H.D. (1995). Age effects in the processing of typical and distinctive faces. *The Quarterly Journal of Experimental Psychology, 48*A(2), 447-465.

MacMillan, N.A., & Creelman, C.D. (1991). *Detection theory: A user's guide*. New York: Cambridge University Press.

Metzger, M.M. (1999). Differential effects of transformation on facial recognition in young children: A pilot study. *Perceptual and Motor Skills, 89*, 799-807.

Podd, J. (1990). The effects of memory load and delay on facial recognition. *Applied Cognitive Psychology, 4*, 47-59.

Shapiro, P.N., and Penrod, S. (1986). Meta-analysis of facial identification studies. *Psychological Bulletin, 100*(2), 139-156.

Sophian, C., and Stigler, J.W. (1981). Does recognition memory improve with age? *Journal of Experimental Child Psychology, 32*, 343-353.

ACKNOWLEDGEMENTS: This research was supported, in part, by a research and development grant awarded by The Pennsylvania State University. I am especially grateful to John Podd for helpful comments and suggestions on an earlier draft of this manuscript. I also thank the teachers at the Mayfield Center School (Mayfield Village, OH) for permission to conduct this experiment. Portions of this paper were presented at the annual meeting of the Midwestern Psychological Association, Chicago, IL, May 2000. Requests for reprints may be sent to Mitchell Metzger, The Pennsylvania State University, Shenango campus, 147 Shenango Ave., Sharon, PA, 16146. (Email: mxm69@psu.edu)

Relationships between Dimensions of Attachment and Empathy

Jeffrey A. Joireman
Tami Lynn Needham
Amy-Lynn Cummings
Seattle Pacific University

Two studies explored the relationship between dimensions of attachment (closeness, trust, anxiety) and three forms of empathy (empathic concern, perspective taking, personal distress). In Study 1, 134 college students completed Collins and Read's (1990) Adult Attachment Scale and Davis' (1983) Interpersonal Reactivity Index. Greater trust and comfort with closeness were associated with greater EC and PT, whereas greater anxiety was associated with greater PD. In Study 2, 261 college students completed the Study 1 measures, along with several additional attachment measures, and a measure assessing interdependent self-construal (Cross, Bacon, & Morris, 2000). Results were generally consistent with those found in Study 1: comfort with closeness predicted greater EC and PT, whereas greater anxiety predicted greater PD. Results also revealed links between an interdependent self-construal, secure attachment dimensions, and the more adaptive forms of empathy, but provided mixed support for the hypothesis that an interdependent self-construal mediates the relationship between attachment and empathy.

Interpersonal relationships function to fulfill human needs of closeness, safety, and security. One of the most influential frameworks for studying functioning in interpersonal relationships has been Bowlby's theory of attachment (Bowlby, 1969, 1973, 1980). It holds that early experiences with one's primary caregiver lay the foundation for functioning in subsequent relationships via the development of internal working models of the self (as worthy of love, or not) and others (as trustworthy or not). Inspired by Hazan and Shaver's (1987) frequently-cited article, many recent studies have focused on how adult attachment styles (and their underlying dimensions) shape functioning in interpersonal relationships. Consistent with Bowlby's (1979, p. 129) claim that attachment processes operate "from the cradle to the grave," a growing number of studies have confirmed that individuals with a secure adult attachment style – characterized by comfort with closeness (low avoidance) and a positive view of self (low anxiety) – enjoy higher

---

Author info: Correspondence should be sent to: Dr. Jeff Joireman, Department of Psychology, Seattle Pacific University, 3307 3rd Avenue West, Seattle, WA 98119. Email: joireman@spu.edu.

North American Journal of Psychology, 2001, Vol. 3, No. 3, 63-80.

© NAJP

quality relationships than those with less secure attachments (e.g., Collins & Read, 1990; Feeney, 1996; Feeney & Noller, 1990, 1991; Hazan & Shaver, 1987; Kobak & Hazan, 1991; Kobak & Sceery, 1988; Levy & Davis, 1988; Simpson, 1990; Simpson, Rholes, & Nelligan, 1992).

These findings raise an important question, namely, why does a secure attachment lead to healthier functioning in interpersonal relationships? The present paper attempts to shed light on this question by examining the relationship between adult attachment and empathy (Davis, 1980, 1983). We hypothesize that a secure attachment style (or dimensions related to a secure style) should be positively related to *empathic concern (EC)* (i.e., tender concerned feelings for others in need) and *perspective taking (PT)* (i.e., the ability to adopt the perspective of another), whereas less secure styles (or dimensions) should be associated with greater *personal distress (PD)* (i.e., feelings of discomfort in emergency situations). We develop this hypothesis by reviewing relevant theory and research on attachment, empathy, and functioning in interpersonal relationships, and then report two preliminary studies on the relationship between attachment and empathy. Study 1 examines the relationship between three dimensions of attachment (comfort with closeness, trust, and anxiety; Collins & Read, 1990) and three forms of empathy (EC, PT, and PD; Davis, 1980, 1983). Study 2 replicates and extends these findings using a broader sampling of categorical and continuous attachment measures (Bartholomew & Horowitz, 1991; Collins & Read, 1990; Fraley, Waller, & Brennan, 2000; Hazan & Shaver, 1987), and exploring whether an *interdependent self-construal* (i.e., the extent to which people think about themselves in terms of their close relationships; Cross, Bacon, & Morris, 2000) mediates the relationship between attachment and empathy. In so doing, the present studies help to extend past studies on adult attachment which, to date, have suggested a link between adult attachment and empathy, but have, to our knowledge, not yet directly tested this relationship (cf. Batson & Shaw, 1991; Schoenrade, Batson, Brandt, & Loud, 1986).

Attachment Theory

While the basics of attachment theory are probably well known to most readers, a brief review of the theory is warranted, given the central role attachment theory accords to the construct of empathy. At its core, attachment theory (Bowlby, 1969, 1973, 1980) assumes that early experiences with one's primary caregiver lay the foundation for functioning in subsequent relationships via the development of internal working models of the self (as a person worthy of love, or not) and others (as trustworthy, or not). Early work with infants established the existence of three attachment "styles" (secure, avoidant, ambivalent) based on the nature of parent-child interactions in the so-called "strange situation" (Ainsworth, Blehar, Waters, & Wall, 1978). These styles were shown to depend in large part on the quality of parental (maternal) caregiving (for a review, see DeWolff & van Ijzendoorn, 1997). The attachment styles, in turn, were shown to influence functioning in peer relationships (for reviews, see Allen & Land, 1999; Thompson, 1999). In terms of the present paper, two findings are particularly noteworthy. First, secure attachments were more likely to develop when the primary caregiver accepted the child (i.e., EC) and was sensitive to the child's needs (i.e., PT). Second, children with a secure attachment style were shown to exhibit greater empathy when compared to children with less secure attachment styles (Denham, 1994; Kestenbaum, Farber, & Sroufe, 1989).

Adult Attachment and Functioning in Interpersonal Relationships

Building on the work of early attachment theorists and researchers, Hazan and Shaver (1987) proposed that attachment styles formed in infancy might very well carry forward into adulthood. Furthermore, adults might reasonably be expected to approach romantic relationships according to one of the three styles originally identified by researchers working with infants (secure, anxious/ambivalent, avoidant). Working within this framework, Hazan and Shaver used the typical descriptions of the three infant attachment styles to develop a three-category measure of attachment, whereby adults select the one attachment style that best describes their approach to close relationships. Hazan and Shaver subsequently demonstrated meaningful links between secure, anxious/ambivalent, and avoidant adult attachment styles, quality of romantic relationships, mental models of relationships, and history of parental care giving. Ever since the publication of Hazan and Shaver's classic article, researchers have been keenly interested in understanding how adult attachment styles (and/or their underlying dimensions) develop, operate, and impact both intrapersonal and interpersonal functioning (for recent reviews, see Bartholomew & Perlman, 1994; Feeney, 1996; Hazan & Shaver, 1994; Simpson & Rholes, 1998). One important and consistent finding in the adult attachment literature is that individuals with a secure adult attachment (or those scoring higher on dimensions related to the secure style) report greater satisfaction in interpersonal relationships (Collins & Read, 1990; Feeney, 1996; Feeney & Noller, 1990, 1991; Hazan & Shaver, 1987; Kobak & Hazan, 1991; Kobak & Sceery, 1988; Levy & Davis, 1988; Simpson, 1990; Simpson et al., 1992). These findings obviously raise the question, why?

One possible explanation, which has received surprisingly little direct attention, is that individuals with a secure attachment style exhibit more empathic concern and perspective taking (Davis, 1983) – which in turn facilitate functioning in interpersonal relationships (Davis & Oathout, 1987, 1992; Franzoi, Davis, & Young, 1985). Similarly, it is possible that individuals with less secure attachments evidence greater personal distress (Davis, 1983) – which in turn negatively impacts functioning in interpersonal relationships (Davis & Oathout, 1987, 1992). Initial support for this hypothesis comes from at least three sources. First, a secure attachment and empathy both imply a willingness to approach others, and both serve to facilitate functioning in interpersonal relationships. Second, as noted earlier, children with a secure attachment style evidence greater levels of empathy, and are more likely to have had parents who modeled empathy. Third, a number of recent studies have demonstrated that adults with a secure adult attachment provide greater support and comfort (i.e., caregiving) to relationship partners (Carnelley, Pietromonaco, & Jaffe, 1996; Feeney, 1996; Fraley & Shaver, 1998; Kunce & Shaver, 1994; Simpson et al., 1992; Simpson, Rholes, & Phillips, 1996) and strangers (Westmaas & Silver, 2001) faced with a variety of stress-provoking situations (e.g., impending separations; difficult discussions; life crises).

While the previous findings suggest a link between adult attachment and empathy, little research has directly examined the relationship between adult attachment and the three forms of empathy noted earlier (i.e., EC, PT, and PD; Davis, 1980, 1983). In one exception, Erlanger (1996), using a sample of college students, found no overall differences between secure, anxious/ambivalent, and avoidant groups on either EC or PT.[1] Erlanger did find, however, a significant relationship between attachment style and PD. Specifically, the anxious/ambivalent group reported higher levels of PD as compared to secure and avoidant groups. Viewed in light of the theory and research reviewed earlier, Erlanger's failure to find strong relationships between attachment and empathy is somewhat puzzling. A closer examination, however, suggests one plausible explanation for these findings, namely, that categorical measures of attachment tend to suffer from a number of measurement problems which can weaken their association with related constructs. Three commonly noted problems with such categorical measures include the fact that (a) individuals are (typically) forced to choose just one (multifaceted) style, despite the fact that various aspects of that style may not accurately reflect their own style; (b) individuals are (typically) not allowed to rate the degree to which any given style characterizes their own style, preventing finer discriminations among individuals with the same style, and (c) single-item indicators are inherently less reliable than multi-item measures (Brennan, Clark, & Shaver, 1998; Fraley & Waller, 1998). Given these problems, the present studies focus (primarily) on the relationship between continuous measures of attachment and empathy.

Assuming that adult attachment is best described in terms of its underlying dimensions, it is appropriate to consider how the two major dimensions (closeness and anxiety) might relate to EC, PT, and PD. As a general hypothesis, we expect that a secure attachment style (and those dimensions related to a secure style; high comfort with closeness, high trust, low anxiety) should be positively related to EC (*Hypothesis 1*) and PT (*Hypothesis 2*), whereas less secure styles (or dimensions) should be associated with greater PD (*Hypothesis 3*). A closer look at the studies on adult attachment and caregiving suggest a qualification to this general hypothesis, namely, that the two attachment dimensions may differ in the strength of their relationship with these three forms of empathy. More specifically, several studies suggest that EC and PT (i.e., indicators of warmth, understanding, and support) should be more closely related to the attachment dimension of closeness, whereas PD (i.e., an indicator of self-reported distress in emergency situations) should be more strongly related to the attachment dimension of anxiety (*Hypothesis 4*) (e.g., Feeney, 1996; Simpson et al., 1992, 1996; Westmaas & Silver, 2001).

As a preliminary test of these hypotheses, we conducted two studies using a convenience sample of college students. Study 1 examined the relationship between Collins and Read's (1990) Adult Attachment Scale (AAS) and Davis' (1980, 1983) measure of empathy. Study 2 replicated and extended Study 1 by including a broader sampling of attachment measures, and testing the additional hypothesis that the relationship between attachment and empathy may be mediated by an interdependent self-construal (*Hypothesis 5*) (Cross et al., 2000). We based this last prediction on two lines of reasoning. First, Cross et al. have demonstrated that individuals scoring high on their measure of interdependent self-construal score higher on EC, and to a lesser extent, PT. Second, according to attachment theory (Bowlby, 1969, 1973, 1980), individual differences in attachment styles reflect varying working models of self (as an individual worthy of love, or not) and others (as dependable and willing to provide love, or not). Within this framework, it seemed reasonable to assume that individuals with a secure style (or those scoring higher on dimensions related to a secure style), who hold positive working models of both self and others, and who feel more comfortable being close to others, would be more likely to define themselves in terms of their relationships. On the other hand, it seemed likely that individuals with less secure styles, especially those with negative working models of others, who feel uncomfortable with closeness, would be less likely to define themselves in terms of their close relationships with others.

STUDY 1

Method

Participants and Procedure. College students (N = 134; 26 Men, 108 Women; Median Age = 19) enrolled in a course on relationship development completed a survey containing all measures in class in exchange for extra credit, and were debriefed at the end of the quarter. The sample was composed of Caucasians (92.5%), Asian Americans (3%), and African Americans (0.7%). The remaining students identified with an unspecified ethnic group (4%). The majority of students (60%) indicated that they were not currently in a dating relationship.

Attachment Dimensions. We assessed attachment using Collins and Read's (1990) *Adult Attachment Scale* (AAS). The AAS contains 18 attachment-related statements, rated on a scale from 1 (disagree strongly) to 7 (agree strongly), the majority of which come from Hazan and Shaver's (1987) three-category Adult Attachment Styles questionnaire. Collins and Read's factor analysis of these statements yielded three factors labeled *depend* (e.g., I find it difficult to allow myself to depend on romantic partners; α = .78: all alphas based on current study), *closeness* (e.g., I feel comfortable sharing my private thoughts and feelings with my partner; α = .63) and *anxiety* (e.g., I worry that romantic partners won't care about me as much as I care about them; α = .65).[2]

Empathy Dimensions. Participants also completed three of the four subscales from Davis' (1980, 1983) *Interpersonal Reactivity Index (IRI)*, including *EC* (α = .79), *PT* (α = .80), and *PD* (α = .78). Each scale contains seven Likert-type items rated on a scale from 1 (never describes me) to 5 (always describes me). For example, one EC item reads "I often have tender concerned feelings for people less fortunate than me"; one PT item reads "I try to look at everybody's side of a disagreement before I make a decision;" and one PD item reads "I sometimes feel helpless when I am in the middle of a very emotional situation."[3] Several studies support the ecological validity of these three forms of empathy with regard to functioning in interpersonal relationships (Davis & Oathout, 1987, 1992; Franzoi et al., 1985).

Results

Data Screening. Data screening revealed no problems with skewness or kurtosis, and no univariate outliers were detected (p < .001, two-tailed). Subsequent inspection of the data supported the assumptions of linearity and homoscedasticity. However, multiple regression analyses revealed a single multivariate outlier within the regression analyses on EC (Studentized Deleted Residual < -3.00), and perspective-taking (SDR > 3.00). Each case was dropped within its respective analyses, and both cases were dropped prior to computing the simple correlations.

Test of Hypotheses. As a first step in our analyses, we computed simple correlations between the dimensions of attachment and empathy (see Table 1). Consistent with Hypotheses 1 and 2, EC and PT were positively correlated with dimensions of attachment reflecting a secure attachment (closeness and depend/trust), and PT was negatively associated with the anxiety attachment dimension. Consistent with Hypothesis 3, PD was positively correlated with the anxiety attachment dimension, and negatively correlated with the closeness attachment dimension.

TABLE 1 Simple Correlations and Regression Models Predicting Three Forms of Empathy: Study 1

	Empathic Concern			Perspective Taking			Personal Distress		
Variables	r	β	t	r	β	t	r	β	t
CR Anxiety	.03	.14	1.56	-.20*	-.11	-1.28	.44**	.43	5.50**
CR Close	.21*	.11	1.11	.21*	.08	0.84	-.27**	-.20	-2.30*
CR Depend	.31**	.24	2.43*	.31**	.25	2.49*	-.14	.06	0.63
Gender	.17+	.15	1.69+	-.00	-.04	-0.42	.22*	.20	2.62**
Summary		Adj-R^2	F		Adj-R^2	F		Adj-R^2	F
Overall Model		.09	4.19**		.09	4.13**		.26	12.65**
Change		.09	4.20**		.12	5.51**		.24	13.99**

Note. Gender (1 = men, 2 = women). N's range from 132 to 133 following removal of outliers. CR = Collins and Read (1990). R^2 change and F change values for attachment dimensions, over and above a regression model with gender as the sole predictor. Because R^2 change does not adjust for the number of model parameters, it can, as shown under PT, exceed the overall Adj-R^2 for the entire model. + $p < .10$, * $p < .05$, ** $p < .01$ (two-tailed).

Given the statistical overlap between the attachment dimensions, and our hypothesis that the attachment dimensions might differ in the strength of their relationship with various forms of empathy, we subsequently conducted three multiple regression analyses to examine each attachment dimension's unique relationship with the three forms of empathy (see Table 1). Because gender showed significant relationships with EC and PD, we included it as a covariate in our analyses.

As shown in Table 1, EC and PT were uniquely predicted by the depend attachment dimension, and women reported marginally higher levels of EC. In contrast, PD was uniquely predicted by the anxiety and closeness dimensions, and women reported higher levels of PD. In sum, the findings provided reasonable support for Hypothesis 4, which predicted that the closeness and/or depend attachment dimensions should be more closely associated with the adaptive forms of empathy (EC and PT), whereas the anxiety attachment dimension should be more closely associated with the maladaptive form of empathy (PD).

Discussion

The purpose of the present study was to examine the relationship between three dimensions of attachment identified by Collins and Read (1990) and three forms of empathy identified by Davis (1980, 1983). As predicted, attachment dimensions associated with a secure attachment (closeness and depend/trust) were positively associated with empathic concern (EC) and perspective taking (PT), whereas the anxiety attachment dimension was negatively associated with PT, and positively associated with personal distress (PD). One potentially important result of the present study was the finding that the "depend" dimension was the only unique predictor of both EC and PT. One possible explanation for this finding is simply that the depend dimension showed a higher reliability when compared to the closeness and anxiety dimensions (see Table 1). Another possible explanation is that EC and PT are, indeed, more closely associated with the depend dimension. If true, the latter interpretation seems to suggest the importance of retaining a distinction between the depend and closeness dimensions, rather than combining them into a broader "approach/avoidance" dimension.

To test the stability and generalizability of our initial findings, we conducted a second study in which we incorporated a number of additional categorical and continuous measures of attachment. Based on past research, we anticipated that the continuous measures – where individuals are defined by a latent variable measured by the scale items they rate – would be more closely associated with the various forms of empathy than would the categorical measures of attachment – where individuals pick the attachment style that best describes them. As noted in the introduction, we also hypothesized that the relationship between attachment and empathy might be mediated by an interdependent self-construal Cross et al., 2000).

STUDY 2

Method

Participants and Procedure. College students (*N* = 261; 88 Men, 172 Women, 1 Unidentified; Median Age = 19) enrolled in an introductory psychology course completed a survey containing all measures in class for extra credit. The sample was composed of Caucasians (85.4%), Asian Americans (5.0%), African Americans (1.1%), Hispanics (2.3%), and Native Americans (0.8%). The remaining students identified with an unspecified ethnic group (5.4%). As before, the majority of students (65.5%) indicated that they were not currently in a dating relationship.

Attachment. To examine the replicability and generalizability of the findings from Study 1, we assessed attachment using a variety of measures. As in Study 1, participants completed Collins and Read's (1990) AAS (i.e., depend, closeness, and anxiety. Participants also completed Fraley, Waller and Brennan's (2000) *Experiences in Close Relationships – Revised* scale (ECR-R), an updated version of the *Experiences in Close Relationships* scale (Brennan et al., 1988) based on item response theory, which assesses two major attachment dimensions including *avoidance* ($\alpha = .93$) and *anxiety* ($\alpha = .87$). Each subscale is comprised of 18 attachment-related statements, which participants rate on a scale from 1 (strongly disagree) to 7 (strongly agree). The majority of the items come from Brennan et al.'s original ECR scale.

Participants also completed two categorical measures of attachment. First, participants completed Hazan and Shaver's (1987) *Adult Attachment Styles* questionnaire where individuals select one of three attachment styles that best describes them: *secure, avoidant, anxious/ambivalent.* Second, participants completed Bartholomew and Horowitz's (1991) *Relationship Questionnaire,* based on Bowlby's notion of working models of self and others. Here, individuals select one of four styles that best describes them: *secure* (positive working models of self and others), *fearful-avoidant* (negative models of self and others), *preoccupied* (negative self model, positive other model) and *dismissing-avoidant* (positive self model, negative other model). The first three categories on the RQ correspond to Hazan and Shaver's secure, avoidant, and anxious/ambivalent categories, whereas the fourth style represents a new avoidant classification.

Empathy. As before, participants also completed *EC* ($\alpha = .78$), *PT* ($\alpha = .78$), and *PD* ($\alpha = .74$) subscales from Davis' (1983) *IRI* (see Study 1).

Interdependent Self-Construal. Finally, participants completed Cross et al.'s (2000) *Relational-Interdependent Self-Construal Scale* (RISC). The RISC assesses the extent to which people think about themselves in terms of their close relationships. Respondents indicate their level of agreement with 11 items using a scale from 1 (strongly disagree) to 7 (strongly agree) (e.g., In general, my close relationships are an important part of my self-image). The scale demonstrates high internal reliability, and good convergent and discriminant validity ($\alpha = .86$).

Results

Attachment Dimensions and Empathy. We begin by discussing the relationship between the dimensional attachment measures and the three forms of empathy. Data screening on the continuous measures revealed no problems with skewness or kurtosis, but one univariate outlier was detected on EC ($p < .001$, two-tailed). Subsequent inspection of the data supported the assumptions of linearity and homoscedasticity. Multiple regression analyses revealed a total of four multivariate outliers (SDR > ± |3.00|), one of which corresponded to the univariate outlier. Given that they constituted a small percentage of the cases, these observations were dropped within their respective analyses, and all four cases were dropped prior to computing the simple correlations between all variables. Table 2 shows the simple correlations and multiple regressions bearing on the relationship between the dimensions of attachment and empathy.

TABLE 2 Simple Correlations and Regression Models Predicting Three Forms of Empathy: Study 2

	Empathic Concern			Perspective Taking			Personal Distress		
Variables	r	β	t	r	β	t	r	β	t
Model 1									
CR Anxiety	-.20**	-.08	-1.27	-.08	.01	0.10	.29**	.30	4.55**
CR Close	.30**	.21	2.84**	.15*	.05	0.62	-.24**	-.16	-2.10*
CR Depend	.26**	.06	0.77	.17*	.07	0.76	-.20**	.02	0.29
Gender	.40**	.37	6.47**	.17*	.18	2.79**	.26**	.29	4.83**
Model 2									
FWB Anxiety	-.24**	-.09	-1.27	-.07	.05	0.72	.30**	.23	3.28**
FWB Avoid	-.31**	-.22	-3.21**	-.17*	-.18	-2.39*	.23**	.14	2.11*
Gender	.40**	.15	1.69+	.17*	.18	2.78**	.26**	.29	4.85**

Summary	Adj-R^2	F	Adj-R^2	F	Adj-R^2	F
Model 1						
Overall Model	.23	18.82**	.03	2.84*	.19	14.74**
Change	.09	9.02**	.01	0.92	.14	13.11**
Model 2						
Overall Model	.21	20.68**	.05	4.97**	.16	15.34**
Change	.07	10.40**	.02	3.00+	.10	13.90**

Note. Gender (1 = men, 2 = women). N's range from 229 to 238, following removal of outliers (see text). following removal of outliers (see text). CR = Collins and Read (1990); FWB = Fraley, Waller, and Brennan (2000). R^2 change and F change values for attachment dimensions, over and above a regression model with gender as the sole predictor. + $p < .10$, * $p < .05$, ** $p < .01$ (two-tailed).

While the results did not provide a perfect replication of the findings reported in Study 1, the overall pattern of correlations between the various attachment dimensions and the three forms of empathy was generally consistent with predictions. Consistent with Hypothesis 1, greater EC was associated with higher scores on Collins and Read's

(1990) closeness and depend attachment dimensions, lower scores on Fraley et al.'s (2000) anxiety and avoidance dimensions, and lower scores on Collins and Read's anxiety dimension. In partial support of Hypothesis 2, scores on PT showed similar associations with the attachment dimensions, but the correlations were weaker, and the anxiety attachment dimensions were unrelated to PT. In support of Hypothesis 3, PD showed positive correlations with the anxiety attachment dimensions from both scales, and somewhat weaker correlations with the avoid, closeness, and depend attachment dimensions.

In order to assess the unique relationship between the various attachment dimensions and the three forms of empathy, we next conducted a series of multiple regression analyses. To avoid problems of multicollinearity introduced by the strong theoretical and empirical overlap between the two attachment scales used in Study 2, we did not include the two attachment inventories in the same regression analysis. Rather, as shown in Table 2, we conducted six multiple regression analyses (3 measures of empathy x 2 attachment inventories), using gender as a covariate in each analysis.

We begin by discussing the relationship between empathy and the attachment dimensions from Collins and Read's (1990) scale (top half of Table 2). In contrast to Study 1, the depend dimension was not a unique predictor of EC or PT, but the closeness dimension was a unique predictor of EC, in partial support of Hypothesis 1 and 4. Consistent with Study 1, the anxiety attachment dimension was the best predictor of PD, in support of Hypotheses 3 and 4, but closeness also showed a unique (but weaker, and negative) relationship with PD (Hypothesis 1).

As shown in the bottom half of Table 2, the three forms of empathy showed similar relationships with the attachment dimensions from Fraley et al.'s (2000) inventory. EC was uniquely predicted by (a lack of) avoidance (in support of Hypothesis 1 and 4), and PD was best predicted by the anxiety attachment dimension (Hypothesis 3 and 4), and to a lesser extent, the avoidance dimension (Hypothesis 3). One difference between the results from the two attachment inventories is that the Fraley et al. dimension of closeness was a unique predictor of PT (Hypothesis 2), whereas the Collins and Read dimension of closeness was not.

Mediation Analyses. A second goal of Study 2 was to test the hypothesis that an interdependent self-construal would mediate the relationship between attachment and empathy (Hypothesis 5). For the mediation model to receive support, several conditions must be met: the predictor variable (attachment dimensions) should relate to (a) the criterion variable (empathy) and (b) the potential mediator (interdependent self-construal); (c) the potential mediator should relate to the criterion variable; and (d) the relationship between the predictor variable and the criterion variable should be reduced to non-significant levels (full mediation), or reduced in strength (partial mediation) after statistically controlling for the potential mediator (cf. Baron & Kenny, 1986).

In line with Hypothesis 5, scores on Cross et al.'s interdependent self-construal scale correlated in the anticipated direction with the attachment dimensions of closeness ($r = .24$, $p < .01$) and trust ($r = .25$, $p < .01$) from Collins and Read's (1990) scale, avoidance ($r = -.25$, $p < .01$) from Fraley et al.'s (2000) scale, and the two adaptive forms of empathy, EC ($r = .27$, $p < .01$) and PT ($r = .16$, $p < .05$). Combined with the earlier results linking the various attachment dimensions with various forms of empathy (mediation condition (a)), these results provide additional support for the hypothesis that an interdependent self-construal may mediate the relationship between attachment and empathy (mediation conditions (b) and (c) noted above). To test the final mediation condition, we conducted a number of multiple regression analyses in which we examined the relationship between the various attachment dimensions and EC and PT, after controlling for an interdependent self-construal (because PD was not related to an interdependent self-construal, the mediation model for PD is already rejected). The results of these analyses did not provide strong support for the mediation hypothesis: while the relationship between attachment dimensions (close and lack of avoidance) and EC and PT was reduced after controlling for an interdependent self-construal, the reverse was equally true. As such, the present results do not allow a clear statement regarding the temporal ordering of these various constructs. Nevertheless, the meaningful overlap between attachment, interdependent self-construal, and empathy points to the potential value of future research in this area.

Attachment Categories. For purposes of comparison, we also analyzed the relationship between empathy and two categorical measures of attachment (Bartholomew & Horowitz, 1991; Hazan & Shaver, 1987). For each attachment measure, we conducted a multivariate analysis of variance (MANOVA) on the three forms of empathy, using attachment category and gender as predictor variables. Preliminary screening of the data revealed no problems with skewness or kurtosis. Additional screening supported the assumptions of linearity and homogeneity of variance, and revealed no univariate outliers. However, one multivariate outlier (Mahalonobis Distance = 16.33, $p < .001$) was identified using the Bartholomew and Horowitz (1991) measure, and was thus deleted.

The first analysis failed to reveal a significant relationship between Hazan and Shaver's (1987) three category attachment measure and the three forms of empathy at the multivariate level, Wilk's Lambda = .96, $F_{(6,\ 434)} = 1.51$, $p = .17$.[4] The second analysis did, however, reveal a significant relationship between Bartholomew and Horowitz's (1991) four category attachment measure the three forms of empathy at the multivariate level, Wilk's Lambda = .87, $F(_{9,\ 509}) = 3.28$, $p < .001$. Univariate analyses of variance revealed a significant association between attachment category and EC, $F(_{3,\ 211}) = 2.83$, $p < .05$, and PD, $F(_{3,\ 211}) = 6.89$, $p < .001$ – results for PT were not significant ($F < 1$). Focusing on EC, Tukey tests revealed only one significant difference: individuals reporting a secure style scored significantly higher on EC ($M = 3.95$) than did those reporting a fearful-avoidant style ($M = 3.72$) ($p < .05$). Turning to PD, Tukey tests revealed two significant differences: individuals reporting a secure style scored significantly lower on PD (M = 2.30) than did those reporting either a preoccupied ($M = 2.74$) or a fearful-avoidant style ($M = 2.80$) (*p's* < .001). The effect for gender was significant in both MANOVAs (*p's* < .001), but the gender by attachment interactions were not. MANOVAs using Type I Sums of Squares, which provide more power with unequal cell sizes, yielded identical results.

Discussion

The purpose of Study 2 was twofold. First, we sought to replicate and extend findings from Study 1 using a broader set of attachment measures. Second, we sought to test the hypothesis that an interdependent self-construal (Cross et al., 2000) might mediate the relationship between attachment and empathy. With regard to the former, the present results partially replicated findings from Study 1. Both studies demonstrated that secure attachment dimensions (depend in Study 1, closeness in Study 2) predicted higher EC and PT, whereas the anxiety attachment dimension (in both studies) predicted higher PD. Analyses using categorical attachment measures evidenced weaker results, providing additional support for the superiority of attachment dimensions over attachment categories. We suspect these results reflect the fact that dimensional measures of attachment resolve several of the problems associated with categorical measures of attachment where individuals choose a single, multi-faceted style rather than rating the extent to which they agree with a larger list of attachment-related items (cf. Brennan et al., 1998; Fraley & Waller, 1998). With regard to our second purpose, our results provided limited support for the claim that an interdependent self-construal mediates the relationship between attachment and empathy. Nevertheless, an interdependent self-construal was positively associated with the attachment dimension of closeness, as well as EC and PT.

GENERAL DISCUSSION

Ever since the publication of Hazan and Shaver's (1987) important article, interest in adult attachment has been high. The present studies sought to extend this literature by examining the relationship between attachment and empathy within a college sample. Based on the literature, we predicted that dimensions related to a secure attachment (low anxiety and low avoidance) would facilitate positive forms of empathy (EC and PT), whereas dimensions related to an insecure attachment (high anxiety and avoidance) would be associated with a maladaptive form of empathy (PD). Both studies supported this general hypothesis. Results also revealed some preliminary support for the prediction that secure attachment and positive forms of empathy would be positively associated with an interdependent self-construal (Cross et al., 2000). The present results help to extend past research on adult attachment processes, and may serve as the basis for targeted interventions for individuals experiencing attachment-related difficulties in interpersonal relationships, issues to which we now turn.

As noted in the introduction, many studies have demonstrated links between a secure adult attachment and healthier functioning in interpersonal relationships, and a number of recent studies have also demonstrated that individuals with a secure attachment style exhibit greater caregiving behavior toward partners and strangers faced with a stressful situation. The present studies extended this work by demonstrating links between adult attachment and empathy. More specifically, the present results suggest one possible explanation for the link between adult attachment, interpersonal functioning, and caregiving, namely, that individuals with a secure attachment experience greater EC and PT, which in turn facilitate functioning in interpersonal relationships (e.g., Davis & Oathout, 1987, 1992; Franzoi et al.,1985), and presumably serve as prerequisites for supportive caregiving. Because we did not directly assess interpersonal functioning and/or caregiving in on-going romantic/close relationships, future research will be necessary to clearly test whether empathy serves as a mediator between adult attachment, healthier interpersonal functioning, and greater caregiving.

In addition to identifying one potential mediator in the relationship between attachment and functioning in interpersonal relationships, the present research provides new insight into the underlying basis of adult attachment. Consistent with the notion that adult attachment is based on working models of the self and others, our results suggest that individuals scoring low on avoidance are more likely to report an interdependent self-construal (Cross et al., 2000). These results, while preliminary, would seem to identify another potential mediator in the relationship between adult attachment and functioning in interpersonal relationships, namely, an interdependent self-construal. Indeed, individuals scoring low on avoidance (and high in an interdependent self-construal) may engage in more constructive relationship behaviors, in part, because such individuals blur the line between self and partner.

On an applied note, the present results might also serve to identify potential points of intervention for individuals experiencing attachment-related difficulties in interpersonal relationships. In particular, our results suggest that individuals with insecure attachments might experience an improvement in their relationships if they can be trained to experience more EC and greater PT. Given the links between attachment, empathy, and an interdependent self-construal, interventions might also focus on encouraging individuals with an insecure attachment to define themselves in more interdependent terms, assuming that future research can clearly demonstrate that an interdependent self-construal does, in fact, influence empathy.

Limitations and Future Directions

While we believe the current studies help shed additional light on adult attachment processes, the present results are most appropriately treated as preliminary, in light of several limitations. First, the correlational nature of the data prevents conclusive statements regarding causality. As such, future longitudinal studies will be necessary to clearly determine whether empathy serves as a mediator in the relationship between adult attachment and functioning in interpersonal relationships, and whether an interdependent self-construal may mediate the relationship between attachment and empathy. Second, the relationships between attachment and empathy, while statistically significant, were not especially large, indicating that much of the variance in empathy is explained by factors other than attachment. This is especially true of PT in Study 2, where the attachment measures and gender explained only three percent of the variance in PT. Finally, given that both studies used convenience samples of (mostly Caucasian) college students, the majority of whom were not currently in a romantic relationship, the present results may be somewhat limited in generalizability. While we see no theoretical reason why such results should fail to replicate within a broader sample, it would nevertheless be wise to replicate these results across a broader sample of adults and types of relationships.

REFERENCES

Ainsworth, M. D. S., Blehar, M. C., Waters, E., & Wall, S. (1978). *Patterns of attachment: A psychological study of the Strange Situation.* Hillsdale, NJ: Erlbaum.

Allen, J. P., & Land, D. (1999). Attachment in adolescence. In J. Cassidy and P. R. Shaver (Eds.), *Handbook of attachment* (pp. 319-335). New York: Guilford Press.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*, 1173-1182.

Bartholomew, K., & Horowitz, L. M. (1991). Attachment styles among young adults: A test of a four-category model. *Journal of Personality and Social Psychology, 61*, 226-244.

Bartholomew, K., & Perlman, D. (Eds.). (1994). *Advances in personal relationships: Vol. 5. Attachment processes in adulthood.* London: Jessica Kingsley.

Batson, C. D., & Shaw, L. L. (1991). Evidence for altruism: Toward a pluralism of prosocial motives. *Psychological Inquiry*, 2, 107-122.

Bowlby, J. (1969/1982). *Attachment and loss. Vol. 1. Attachment.* New York: Basic Books.

Bowlby, J. (1973). *Attachment and loss. Vol. 2. Separation.* New York: Basic Books.

Bowlby, J. (1979). *The making and breaking of affectional bonds.* London: Tavistock.

Bowlby, J. (1980). *Attachment and loss. Vol. 3. Loss.* New York: Basic Books.

Brennan, K. A., Clark, C. L., & Shaver, P. R. (1998). Self-report measurement of adult romantic attachment: An integrative overview. In J. A. Simpson & W. S. Rholes (Eds.), *Attachment theory and close relationships* (pp. 46-76). New York: Guilford Press.

Carnelley, K. B., Pietromonaco, P. R., & Jaffe, K. (1996). Attachment, caregiving, and relationship functioning in couples Effects of self and partner. *Personal Relationships, 3*, 257-278.

Collins, N. L., & Read, S. J. (1990). Adult attachment, working models, and relationship quality in dating couples. *Journal of Personality and Social Psychology, 58*, 644-663.

Cross, S. E., Bacon, P. L., & Morris, M. L. (2000). The relational-interdependent self-construal and relationships. *Journal of Personality and Social Psychology, 78*, 791-808.

Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. JSAS: *Catalog of Selected Documents in Psychology, 10*, 85.

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology, 44*, 113-126.

Davis, M. H., & Oathout, H. A. (1987). Maintenance of satisfaction in romantic relationships: Empathy and relational competence. *Journal of Personality and Social Psychology, 53*, 397-410.

Davis, M. H., & Oathout, H. A. (1992). The effect of dispositional empathy on romantic relationship behaviors: Heterosocial anxiety as a moderating influence. *Personality and Social Psychology Bulletin, 18*, 76-83.

Denham, S. A. (1994). Mother-child emotional communication and preschoolers' security of attachment and dependency. *Journal of Genetic Psychology, 155*, 119-121.

DeWolff, M., & van Ijzendoorn, M. (1997). Sensitivity and attachment: A meta-analysis on parental antecedents of infant attachment. *Child Development, 52*, 857-865.

Erlanger, D. M. (1996). *Empathy, attachment style, and identity development in undergraduates.* Unpublished doctoral dissertation, Columbia University.

Feeney, J. A. (1996). Attachment, caregiving, and marital satisfaction. *Personal Relationships, 3*, 401-416.

Feeney, J. A., & Noller, P. (1990). Attachment style as a predictor of adult romantic relationships. *Journal of Personality and Social Psychology, 58*, 281-291.

Feeney, J. A., & Noller, P. (1991). Attachment style and verbal descriptions of romantic relationships. *Journal of Social and Personal Relationships, 8*, 187-215.

Fraley, R. C., & Shaver, P. R. (1998). Airport separations: A naturalistic study of adult attachment dynamics in separating couples. *Journal of Personality and Social Psychology, 75*, 1198-1212.

Fraley, R. C., & Waller, N. G. (1998). Adult attachment patterns: A test of the typological model. In J. A. Simpson & W. S. Rholes (Eds.), *Attachment theory and close relationships* (pp. 77-114). New York: Guilford Press.

Fraley, R. C., Waller, N. G., & Brennan, K. A. (2000). An item response theory analysis of self-report measures of adult attachment. *Journal of Personality and Social Psychology, 78*, 350-365.

Franzoi, S. L., Davis, M. H., & Young, R. D. (1985). The effects of private self-consciousness and perspective taking on satisfaction in close relationships. *Journal of Personality and Social Psychology, 48*, 1584-1594.

Hazan, C., & Shaver, P. (1987). Romantic love conceptualized as an attachment process. *Journal of Personality and Social Psychology, 52*, 511-524.

Hazen, C., & Shaver, P. R. (1994). Attachment as an organizational framework for research on close relationships. *Psychological Inquiry, 5*, 1-22.

Kestenbaum, R., Farber, E. A., & Sroufe, L. A. (1989). Individual differences in empathy among preschoolers: Relation to attachment history. In N. Eisenberg (Ed.) *New directions in child development: No. 44. Empathy and related emotional responses* (pp. 51-64). San Francisco: Jossey-Bass.

Kobak, R. R., & Hazan, C. (1991). Attachment in marriage: Effects of security and accuracy of working models. *Journal of Personality and Social Psychology, 60*, 861-869.

Kobak, R. R., & Sceery, A. (1988). Attachment in late adolescence: Working models, affect regulation and representations of self and others. *Child Development, 59*, 135-146.

Kunce, L. J., & Shaver, P. R. (1994). An attachment-theoretical approach to caregiving in romantic relationships. In K. Bartholomew & D. Perlman (Eds.), *Advances in personal relationships: Vol. 5. Attachment processes in adulthood* (pp. 205-237). London: Jessica Kingsley.

Levy, M. B., & Davis, K. E. (1988). Love styles and attachment styles compared: Their relations to each other and to various relationship characteristics. *Journal of Social and Personal Relationships, 5*, 439-471.

Schoenrade, P. A., Batson, C. D., Brandt, J. R., & Loud, R. E. (1986). Attachment, accountability, and motivation to benefit another not in distress. *Journal of Personality and Social Psychology, 51*, 557-563.

Simpson, J. A. (1990). Influence of attachment styles on romantic relationships. *Journal of Personality and Social Psychology, 59*, 971-980.

Simpson, J. A., & Rholes, W. S. (Eds.). (1998). *Attachment theory and close relationships.* New York: The Guilford Press.

Simpson, J. A., Rholes, W. S., & Nelligan, J. S. (1992). Support seeking and support giving within couples in an anxiety-provoking situation: The role of attachment styles. *Journal of Personality and Social Psychology, 62*, 434-446.

Simpson, J. A., Rholes, W. S., & Phillips, D. (1996). Conflict in close relationships: An attachment perspective. *Journal of Personality and Social Psychology, 71*, 899-914.

Thompson, R. A. (1999). Early attachment and later development. In J. Cassidy and P. R. Shaver (Eds.), *Handbook of attachment* (pp. 265-286). New York: Guilford Press.

Westmaas, J. L., & Silver, R. C. (2001). The role of attachment in responses to victims of life crises. *Journal of Personality and Social Psychology, 80*, 425-438.

Footnotes

[1] It should be noted, however, that planned comparisons on a sub-sample of sophomores and juniors revealed that avoidant group scored lower on EC, relative to the anxious/ambivalent group.

[2] Several studies suggest that attachment is best described in terms of two underlying dimensions (i.e., anxiety and avoidance; Brennan, Clark, & Shaver, 1998). For purposes of direct comparison, and because the depend and closeness dimensions might show meaningful differences, we retained Collins and Read's (1990) three dimensions.

[3] In Study 1, one PD item was inadvertently omitted. Nevertheless, the consistency in findings for PD across Studies 1 and 2 suggest that the current (six-item) scale is representative of the original (seven-item) scale

[4] Univariate tests indicated a significant relationship between attachment category and PD ($p = .05$), but Tukey post hoc tests failed to reveal any significant differences between the three groups. As such, these results are not discussed further.

Authors' Note: Preparation of this manuscript was supported by a Faculty Research Grant from Seattle Pacific University to the first author. The authors would like to thank Jon Anderson and John Thurston for their help with the data collection, and three anonymous reviewers for the helpful comments on an earlier draft of this paper.

PROGRAMS FOR YOUNG FATHERS: ESSENTIAL COMPONENTS AND EVALUATION ISSUES

Peggy B. Smith
Baylor College of Medicine

Ruth S. Buzi
Baylor College of Medicine

Maxine L. Weinman
University of Houston

The issues surrounding teenage pregnancy and the increase in the number of female-headed households have led to an increased interest in the roles fathers play in shaping their children's lives. Programs targeting young fathers have been identified as one vehicle to increase fathers' responsibility for their children and decrease subsequent pregnancies. These programs must integrate essential components addressing the unique needs of this group. It is also essential that these programs include evaluation strategies that will document whether interventions bring about desired change. In this article, we describe these essential components and the evaluation issues that emerge.

Twenty-four percent of families are headed by females today, as opposed to only 6% in 1960 (U.S. Bureau of the Census, 1998). Children who live in female-headed households face many problems not experienced by children in intact families. The absence of fathers contributes to high poverty rates faced by children living in these families. Moreover, there is some evidence indicating that children with absent fathers are at increased risk for behavioral difficulties, social problems such as youth violence, domestic violence, poor academic performance, and teenage pregnancy (Argys, Peters, Brooks-Gunn, & Smith, 1998). Paternal help with childcare and financial support affects children directly through the father-child relationship and indirectly through paternal influence on maternal behavior (Rhein, Ginsburg, Schwarz, et al., 1997).

The problem of absent fathers is especially significant for children of teenage fathers. About 2% of young men become fathers during

Author info: Correspondence should be sent to: Ruth Buzi, Population Program, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030.

North American Journal of Psychology, 2002, Vol. 4, No. 1, 81-92.
© NAJP

adolescence, with higher rates among inner-city and minority youth (Stouthamer-Loeber & Wei, 1998). Their long-term commitment to parenthood and to the relationship with the child's mother may not be present (Cox & Bithoney, 1995). Although recent US reports suggest that the teenage birthrate has declined between 1991 and 1996 from about 519,600 annually to approximately 494,300 during the 5-year period (Ventura, Curtin, & Mathews, 1998), teenage pregnancy continues to be a major public health issue. Adolescent pregnancy and parenthood result in a cycle of disadvantage for the young woman and her child. School dropout, which often results from early pregnancy, can lead to long-term poverty. Also, the children of adolescent mothers show a higher incidence of developmental delays which persist into adolescence (Roye & Balk, 1996).

Over the past years there has been increasing interest in the roles fathers play in shaping their children's lives. The renewed emphasis on family values and the well being of families in the public policy debate has spotlighted fathers' roles in childbearing. Moreover, the 1996 welfare reform legislation that was written in response to the high rates of teenage pregnancies and female-headed households has emphasized the roles unwed fathers have in being responsible parents (Rangarajan & Gleason, 1998).

Prevention of unplanned teen pregnancies have been endorsed since the 1970s. However, there is a lack of systematic information about the males' role in this endeavor (Guagliardo, Zhihuan, & D'Angelo, 1999; Rhein, Ginsburg, Schwarz, et al., 1997). Little is known about how to reach them or how to influence their behavior (Sonenstein, Stewart, Lindberg, Pernas, & Williams, 1997). Targeting young males is important because they are involved in teen pregnancy, and need ready access to services that promote positive and responsible behaviors (Armstrong, Cohall, Vaughan, et al., 1999).

The purpose of this paper is to suggest essential components in a program targeting young fathers and the evaluation issues that emerge in such programs. While these issues have varying relevance to all programs targeting fathers, the focus of this paper is on those interventions that target men who father children of teen mothers.

Teenage fatherhood is associated with complex behavioral issues that affect teen fathers' involvement with their children. Review of studies on young fathers have suggested that fathering a child as a teenager is correlated with high risk behavior such as drug abuse, sexually transmitted diseases, violence, and school dropout (Spingarn & DuRant, 1996). For example, a study of 399 inner-city African American males ages 12-18 found that a total of 24.2% of males were involved in pregnancy. These males were about 14 times more likely to report three or more sex partners in the last year, more than five times as likely to report a sexually transmitted disease history, more than three times as likely to test positive for drugs, and more than 2.5 times as likely to be inconsistent or nonusers of condoms as compared with males without a pregnancy history (Guagliardo, Zhihuan, & D'Angelo, 1999). Another study of 125 incarcerated adolescents found that 25.6% of the respondents reported delinquency and adolescent fatherhood, and 40.6% of fathers reported having caused more than one pregnancy (Nesmith, Kerman, Oh, & Feinstein, 1997). In the longitudinal Pittsburgh Youth Study, the pregnancy rate among 506 inner city adolescent males from a public school was 12.3% (Stouthamer-Loeber & Wei, 1998). These young fathers were twice as likely to be classified as serious delinquents. Being delinquent and a young father were both related to early sexual experience, drug exposure, cruelty and antisocial behavior. The Youth Risk Behavior Survey of 4,159 students in grades 9-12 in 59 randomly selected public high schools in Massachusetts indicated that of the 824 sexually active males, 12% had been involved in a pregnancy. Furthermore, males who were involved in a pregnancy had other problem behaviors such as substance use, violence, multiple sex partners, and fighting. Data from the National Longitudinal Survey of Youth (NLSY) (Nock, 1998) and data from the National Longitudinal Survey of Labor Market Experience-Youth Cohort (NLSY) (Pirog-Good, 1996) found that teen fathers were more likely to live with single parents in impoverished backgrounds. They were also more likely to become involved in drugs and crime at earlier ages, and were more likely to experience school failure and low wages. Studies also found that over 20% of incarcerated offenders were teenage fathers (McLaughlin, Reiner, Reams & Joost, 1999; Nesmith, Klerman, Oh & Feinstein, 1997).

Studies identified role behaviors that influence the involvement of these young fathers in their children's life. For example, a study among 307 Latino adolescent and adult men, of which 207 impregnated one or more teens, found that their own parental background and male role models mediated their involvement with their children (Goodyear, Newcomb, & Allison, 2000). They hypothesized that constructs such as childhood experiences among males are related to teenage pregnancy as well as gender role beliefs. Additionally, Furstenberg and Weiss (2000) found in a study among 110 males from the Baltimore Parenthood Study that a strong stable presence of biological fathers in their histories affected the timing of their own family formation. Furthermore, if their own fathers lived with them there was a stronger likelihood of living with their own children. The stable presence of a father in their own home is a positive developmental milestone in which fathers are models and mentors for societal expectations and transmission of values from one generation to the next. Christmon (1990) also found in a group of 43 unwed fathers that their willingness to assume parental support for their children was related to their own self-image and role expectations. Therefore, programs for young fathers have attempted to address these complex behavioral and social issues.

Programs for young fathers

Although programs to address the needs of young fathers have been initiated, young fathers have not received the types of services that unwed mothers have (Rangarajan & Gleason, 1998). Most programs are based on four assumptions: (1) if the programs are offered, the fathers will attend. (2) The services will meet the needs of the participants. (3) Once young men enroll in a program they will use the services that are available, and (4) the services will be beneficial for recipients and will produce positive outcomes (Lane & Clay, 2000). Most demonstration projects attempted to address positive outcomes by providing jobs or job training, sex education, counseling and peer support, incentives for children support and encouragement for fathers to take an interest in their children. The importance of employment issues has been emphasized in most programs because of the relationship between earnings, involvement with children and the likelihood of marriage. Kost (1997) examined the results of educational employment and psychosocial support for young fathers and found that "completers" of the program, less than 50% of the sample, tended to have more schooling, be previously employed and participate in program activities. He noted that young fathers' need for emotional support usually exceeds what programs could offer and that staff noted feelings of depression and hopelessness among fathers. Many current researchers also point out that programs tend to emphasize one or two areas such as employment or sex education (Lane and Clay, 2000) or focus research on a few variables without a theoretical framework.

Focus groups have been conducted to identify expectations as well as familial issues among young fathers. For example, Allen and Doherty (1996) in their study of the perceptions of 10 young fathers found that these males are difficult to reach, are guarded in their responses and are hard to locate. When interviewed, they expressed concern about their children but also acknowledged that they themselves were too young to father children and they had insufficient role models to help them make the transition to adulthood. They cited the difficulty of seeing their children because of a poor relationship with the baby's mothers. Dallas and Chen's (1998) school-based focus group study among 5 low-income

African American adolescent fathers identified a number of barriers among adolescent fathers. The major themes that emerged were issues regarding fatherhood transitions, competencies of fathers, role-set relationships, social norms of fatherhood, and father-child contact. Major difficulties included limited sources of income, lack of parenting experiences, abandonment by their own fathers and fatalistic views. Adolescents considered paternity an important role but did not feel ready to fulfill it. Many did not have the negotiation skills to deal with the birthmothers who denied access to their children.

Programs that target young fathers have to define dimensions of father involvement in designing interventions. Several dimensions have been identified in the literature. For example, Marsiglio and Day (1998) suggest that fatherhood includes roles such as nurturance and provision of care, moral and ethical guidance, emotional, practical, and psychosocial support for female partners, and economic provisioning or bread-winning. Levine and Pitt (1995) integrate the concept of responsibility in their definition, and suggest that a responsible father is someone who 1) waits to have a baby until he is prepared emotionally and financially to support his child; 2) establishes his legal paternity if and when he does have a baby; 3) actively shares with the child's mother in the continuing emotional and physical care of their child from pregnancy onwards; and 4) shares with the child's mother in the continuing financial support of their child, from pregnancy onward.

Based on this review, it appears that there are four essential components that need to be integrated in a program for young fathers. These components also include related outcomes that can be measured to document success.

PROPOSED PROGRAM COMPONENTS

1. The legal aspect of fatherhood:

This component includes establishing paternity and child support provision. Programs need to encourage it because there is some evidence that receipt of child support has a positive impact on children's cognitive test scores beyond its contribution to total income (Argys, Peters, Brooks-Gunn, & Smith, 1998). Programs can measure program impact by determining if participants established paternity and provided child support subsequent to program participation. In this area, differences may appear depending on whether or not that support was voluntary or was court-ordered. Some literature suggests that support, in general, varies by marital status and that never married couples, subsequent to the dissolution of the relationship, receive less support for the children (Horn & Bush, 1997). Other studies suggest that age of the child impacts support, and the older the child the less financial and emotional support is provided by the biological father (Rangarajan & Gleason, 1998). Other possible factors include status of the couple's relationship, number of other children, and the presence of another male in the household. Functional support such as the provision of material objects for both the mother and child, sometimes referred to by Rangarajan and Gleason (1998) as "gaming" can also be quantified. This behavior is seen by some as a type of bartering system that provides the father access to his child. Documenting the establishment of paternity and provision of child support or other material contributions can determine program impact.

2. Self-sufficiency:

Self-sufficiency of fathers is essential in order to help these young fathers to be able to financially support their child or pay child support. Programs have to ensure that males become self-sufficient by obtaining adequate employment and also have access to educational programs that provide employment mobility. Self-sufficiency can be measured by job acquisition as well as by factors such as the reduction of age-grade discrepancy for the father, school re-enrollment, and subsequent diploma acquisition. Programs can also measure participation in eligibility establishment for services such as Special Supplemental Nutrition Program for Women, Infants and Children (WIC), Children's Health Insurance Program (CHIPS), and Medicaid. Compliance with immunizations of children can be measured as well.

The procurement and duration of employment as a consequence of program participation and changes in household income can be quantified and be used as an indicator of success. One popular belief suggests that fathers are more likely to spend time with their children and support them financially if they are gainfully employed and have jobs that offer opportunities for advancement and promotion.

3. Effective parenting skills:

The development of effective parenting skills is essential to enhance responsible fatherhood. This supports the hypothesis that fathers can influence the cognitive and personal development of their children (Argys, Peters, Brooks-Gunn, & Smith, 1998; Johnson, 1995). The evaluation process can measure the impact of program participation on the development and well being of the child. Programs could quantify time fathers spend with their children and assess if its quantity can bring about differences in behavior. The variety of activities associated with father-child activities can be quantified as well and categorized as outcome measures. Moreover, the father's acquisition of child development knowledge and child-rearing skills through program participation can also be measured as an outcome. Intermediate determinants of program outcomes can be measured as well. This will include male support of female partners by participating in prenatal care, Lamaze classes, and maternity education programs. Support can be defined as the provision of transportation, acting as a Lamaze coach, or accompanying prenatal care visits. Such variables are observable behavior and can measure program effectiveness.

4. Behavioral Issues

This component refers to the development of skills and behavior promoting a healthy life style among program participants. Since studies suggest that young father engage in high-risk behavior, targeting these behavior appears to be relevant. Selected outcomes for male involvement programs can include reduction of rates of subsequent pregnancies, sexually transmitted infections (STI), violence, and substance use.

Studies among young adults also suggest that there is an association between sexual risk taking behavior and specific mental health problems such as depression, substance abuse, and conduct disorder (Ramrakha, Caspi, Dickson, Moffitt & Paul, 2000). For example, a study in a clinic serving young males found that clients utilized health services in order to also receive mental health attention. While 80% of clients received medical counseling, 20% received mental health counseling (Armstrong, Cohall, Vaughan, et. al., 1999). Therefore, these issues will require interventions in this population. There are a variety of standardized assessment tools that can document rate reduction of symptoms among participants.

ISSUES IN PROGRAM EVALUATION

Participant selection

Two important issues related to participant selection that affect program outcomes are participants' characteristics and age. Since many young fathers have a variety of risks such as criminal records, lack of education, lack of job skills, and children by multiple women (Nesmith, Klerman, Oh, & Feinstein,1997), documenting success is difficult. One significant participant characteristic is whether or not the targeted males have been involved in the judicial system or have been in prison. Understandably, this group may possess the greatest need for assistance in order to establish paternity and a productive relationship with their children. Recruitment may be easily facilitated through linkages with prison release programs. However, several important considerations associated with this strategy should be acknowledged before a previously incarcerated group is targeted for these programs. First, a traditional outcome for male programs involves some formal employment. Historically, businesses are reluctant to hire convicted felons. Many job application questions are now structured to elicit information concerning any legal system encounter even if an offense is minor or a conviction overturned. To omit such data is grounds for immediate termination. Post-incarcerated males may be functionally illiterate or lack prerequisite job readiness skills to successfully maintain employment. To complicate matters, while incarcerated, these fathers may have continued to accrue sizable financial child support obligations and back payments. In some states such as Texas, outstanding child support payments still accumulate during prison confinement. Upon release, until his child support financial obligations are met, the father cannot qualify for a driver's license, limiting employment opportunities. Such constraints may force males to turn to more lucrative but illegal means to address this debt. During their confinement the relationship with their child's mother can also become estranged so that the re-establishment of meaningful child-father relationships is difficult. Negative interpersonal relations may in turn dampen any fervor to reconnect with the child, especially if another male has initiated a relationship with the teen mother.

Participant selection is also affected by age. Not all babies born to adolescent girls are fathered by teenagers. The literature suggests that approximately one-third of the male partners of adolescent mothers are adults (Darroch, Landry, & Oslak, 1999; Landry & Forrest, 1995). An analysis of 46,500 teen births in California (Males and Chew 1996) found that two-thirds of fathers were post-school adults and were an average 6.7 years older than junior high mothers. Nevertheless, while the group is not homogeneous, many projects still target fathers 15-24 years of age. Several program design issues should be weighted when a wide range in age categories exists among recruits. A priority perhaps should be given to adolescent participants. First, teenage males, because of their age, may not have been convicted of a serious crime or may have been afforded leniency as minors for any criminal actions. Second, the younger male may still be enrolled in school, which also provides institutional support for program compliance. Affiliation with their family of origin by living in a family unit, albeit loosely defined, may provide lifestyle structure, which facilitate program participation and protect from some of the harsher realities of life. Third, and perhaps most important, the older male by contrast, may be more likely to have fathered children by several women. Multiple family obligations not only drain the male of already limited resources, but also present the evaluation design with a challenge as to which father-child relationship the program should measure.

Program designs must also include inclusion and exclusion criteria even though programs may be limited by pre formed highly selected groups.

Program retention

In order to document program change, young men must be retained in the program for a defined period of time. Responsibilities associated with employment and school may preclude young men from a major time commitment to structured programs. The program retention and project duration prompts the question of how much program time is sufficient to bring about behavioral change. Programs could profit from pilot efforts to develop the definition for minimal participation criteria.

The issues of selection and retention have an impact on the evaluation. This is because they impact sample size and the ability to document significant change. Barnow and Stapleton (1996), in their analysis of male programs, confirmed that programs may identify at-risk males but the number of enrolled males in most existing intervention projects is small. It is known that young males are reluctant to seek help because they equate seeking help with inappropriate masculine behavior (Armstrong, Cohall, Vaughan,et al.,1999). Another significant issue is attrition as a result of individuals lost to follow-up. This can occur when programs lack adequate tracking methodology. Therefore, programs have to develop effective strategies to track participants.

CONCLUSION

There is consensus in the literature that fathers' participation is essential in the rearing of their children. Additionally, the government has encouraged fathers' participation in order to decrease public subsidy for single mothers and their children. In order to maximize the success of programs for young fathers, programs must be able to address needs of young fathers that are based on research and are translated and delivered through a well-planned program. Programs need to have evaluation strategies to assess the impact of their interventions. We are aware that it will be difficult to document significance due to programmatic issues such as availability of fathers, their ages, their criminal history, as well as their willingness and ability to become involved in interventions. However, unless programs can show some demonstrable impact, it may be impossible to defend their implementation or continued funding. Programs can implement an evaluation plan and deal with these issues as they appear.

REFERENCES

Allen, W.D. & Doherty, W.J. (1996). The responsibilities of fatherhood as perceived by African American teenage fathers. Families in Society: *The Journal of Contemporary Human Services, 77*(3), 142-155.340-350.

Argys, L. Peters, H. E., Brooks-Gunn, J., & Smith, J. R. (1998). The impact of child support on cognitive outcomes of young children, *Demography, 35*, 159-173.

Armstrong, B., Cohall, A. T., Vaughan, R. D., Scott, M., Tiezzi, L., & McCarthy, J. F. (1999). Involving men in reproductive health: The young men's clinic. *American Journal of Public Health, 89*, 902-905.

Barnow, B.S., & Stapleton, D.C. (1996). An Evaluability Assessment of Responsible Fatherhood Programs. The Lewin Group, Inc: Ford Foundation.

Cherlin,A.J (1981). *Marriage, divorce and remarriage: Changing patterns in postwar United States*. Washington,DC: The Urban Institute Press.

Christmon, K. (1990). Parental responsibility of African American unwed adolescent fathers. ***Adolescence***, *25* (99), 645-653.

Cox, J. T., & Bithoney, W. G. (1995). Fathers of children born to adolescent mothers. Predictors of contact with their children at 2 years. *Archives of Pediatric Adolescent Medicine, 149*(9), 962-966.

Dallas, C.M. & Chen, S.C. (1998). Experiences of African American adolescent fathers. *Journal of Nursing Research*, *20*(2), 210-222.

Darroch, J.E., Landry, D.J., & Oslak, S. (1999). Age Differences Between Sexual Partners in the United States. *Family Planning Perspectives, 31*,160-167.

Furstenberg, F. F. & Weiss, C. C. (2000). Intergenerational transmission of fathering roles in at risk families. *Marriage & Family Review*, *29* (2/3), 181-201.

Goodyear, R. K., Newcomb, M.D. & Allison, R.D. (2000). Predictors of Latino men's paternity in teen pregnancy: Test of a mediational model of childhood experiences, gender role attitudes, and behaviors. *Journal of Counseling Psychology*, *47*(1), 116-128.

Guagliardo, M. F., Zhihuan, H., & D' Angelo. (1999). Fathering pregnancies: Marking health-risk behaviors in urban adolescents. *Journal of Adolescent Health, 24*, 10-15.

Horn W., & Bush, A. (1997). *Fathers. Marriage, and Welfare Reform*. Hudson Institute: New York.

Johnson, D.J. (1995). *Father Presence Matters: A Review of the Literature. Toward an Ecological Framework of Fathering and Child Outcomes*. National Center on Fathers and Families.

Kost, K.A. (1997). The effects of support on the economic well-being of young fathers. *Families in Society: The Journal of Contemporary Human Services*, *78*(4), 370-382.

Landry, D.J., & Forest, J.D. (1995). How old are U.S. Fathers? *Family Planning Perspectives, 27*, 159-161.

Lane, T.S. & Clay, C.M. (2000). Meeting the service needs of young fathers. *Child and Adolescent Social Work Journal*, *17*(1), 35-54.

Levine, J.A., & Pitt, E.W. (1995). *New Expectations: Community Strategies for Responsible Fatherhood*. New York, NY.: Family and Work Institute.

Males, M. & Chew, K.S. (1996). The Ages of Fathers in California Adolescent Births. *American Journal of Public Healt,h 86*(4), 565-568.

Marsiglio, W., & Day R. (1998). Social Fatherhood and Paternal involvement: Conceptual, Data, and Policymaking Issues. Nurturing Fatherhood: Improving Data and Research on male Fertility, Family Formation, and Fatherhood.

McLanahan, S. & Sandefur, G.S. (1994). *Growing up With a Single Parent: What Hurts, What Helps*. Cambridge: Harvard University Press.

McLaughlin, C.R., Reiner, S.M. Reams, P. N. & Joost, T.F. (1999). Factors associated with parenting among incarcerated juvenile offenders. *Adolescence, 34*(136), 665-670.

Monthly Vital Statistics Report: Marriage and Divorce Rates United States 1950-94. 43, No. 13, October 23, 1995.

Nesmith, J.D., Klerman, L.V., Oh, M.K., & Feinstein, R.A. (1997). Procreative Experiences and Orientations Toward Paternity Held by Incarcerated Adolescent Males. *Journal of Adolescent Health, 20* (3), 198-203.

Nock, S. L. (1998). The consequences of premarital fatherhood. *American Sociological Review, 63*(2), 250-263.

Pirog-Good, M. (1996). The education and labor market outcomes of adolescent fathers. *Youth & Society, 28*(2), 236-262.

Ramrakha, S., Caspi, A., Dickson, N., Moffitt, T. E, & Paul, C. (2000). Psychiatric disorders and risky sexual behavior in young adulthood: cross sectional study in birth cohort. BMJ , *321*, 263-266.

Rangarajan, A., & Gleason, P. (1998). Young Unwed Fathers of AFDC Children: Do They Provide Support? *Demography, 35*(20), 175-186.

Rhein, L.M., Ginsburg, K. R., Schwarz, D. F., Pinto-Martin, J. A., Zhao, H., Morgan, A.P., & Slap, G.B. (1997). Teen father participation in child rearing: Family perspectives, *Journal of Adolescent Health, 21*, 244-252.

Roye C. F., & Balk S.J. (1996). The relationship of partner support to outcomes for teenage mothers and their children: A review. *Journal of Adolescent Health, 19*, 86-93.

Sonenestein, F.L., Stewart, K., Lindberg, L.D., Pernas, M., & Williams, S. (1997). *Involving Males in Preventing Teen Pregnancy-A Guide for Program Planners*. The California Wellness Institute: The urban Institute.

Spingarn, R.W., & DuRant, R.H. (1996). Adolescents males involved in pregnancy: associated health risk and problem behaviors. *Pediatrics, 98*, (2 Pt 1), 262-268.

Stouthamer-Loeber, M. & Wei, E.H. (1998). The precursors of young fatherhood and its effect on delinquency of teenage males. *Journal of Adolescent Health*, 22(6), 56-65.

Ventura, S.J., Curtin, S.C., & Mathews, T.J. (1998). Teenage Births in the United States: National and State Trends, 1990-96. *Monthly Vital Statistics Report, 46*:11(S) 2: 1-10.

Note: This project was funded by the Texas Department of Health and the Office of Population Affairs and reflects the experiences associated with a male involvement project targeting at-risk minority youth 15-25 years of age.

Meaning in Life and Positive and Negative Well-being

Evelyn D. Scannell, Felicity C. L. Allen & Janet Burton
Monash University

The relationship between meaning in life and well-being was examined by asking adults ($N = 83$) to complete measures of well-being and the revised Life Regard Index (Debats, 1998) that contains affective (Fulfillment) and cognitive (Framework) subscales of meaning in life. Although there were no age differences on Fulfillment, the younger group had significantly lower scores on Framework ($F_{(2, 77)} = 3.98$, $p < .05$) than the older age group. Two hierarchical regression analyses revealed that positive well-being improved prediction of Framework and Fulfillment beyond that afforded by negative well-being measures. One negative (Depression) and two positive (Happiness, Spiritual) well-being measures significantly predicted Framework. No negative and three positive (Happiness, Spiritual, Self-esteem) well-being measures significantly predicted Fulfillment suggesting that affective meaning in life may relate to positive well-being more than it does to negative well-being. Comparison of the two regressions shows that well-being measures were more strongly related to affective meaning (Fulfillment) than to cognitive meaning (Framework). This suggests that, although cognitive and affective meaning are associated with a person's well-being, it is more important to feel that one has meaning in life than to have a structure for that meaning. The association between self-esteem and affective, but not cognitive, meaning indicates that self-esteem is not invariably associated with meaning.

Until recently, the construct of meaning in life had received scant attention in psychological empirical literature (Debats, 1999). Meaning refers to making sense of existence and purpose refers to choosing goals on the basis of an attained sense of order (Reker, Peacock & Wong, 1987; Yalom, 1980). As a broad construct, a sense of meaning can be described as a spiritual dimension (Waisberg & Porter, 1994). Theoretical (Maddi, 1967) and empirical work (Debats, 1990) suggests that meaning in life has cognitive, affective, and behavioural components and an extreme lack of meaning has an affective tone of apathy and boredom that is associated with an absence of actions. People who have a high sense of meaning in life have personally meaningful beliefs and

---

Author info: Correspondence should be sent to: Evelyn Scannell, Psychology Department, Monash University, P.O. Box 197, Caulfield East, Victoria, Australia 3145. E-mail: evelyn.scannell@med.monash.edu.au.

North American Journal of Psychology, 2002, Vol. 4, No. 1, 93-112.

© NAJP

related goals (cognitive component) and feelings and a subjective experience that life is meaningful (affective component) (Antonovsky, 1996; Reker et al., 1987; Shapiro, 1988; Shek, 1991).

Numerous studies suggest that perceiving a meaning in life is related to positive mental health outcomes while meaninglessness is associated with pathology (Antonovsky, 1979, 1987; Debats, 1996; Lindeman & Verkasalo, 1996; Moore, 1997; Reker, 1994; Reker et al., 1987; Ruffin, 1984; Shek, 1991; Zika & Chamberlain, 1992). The main view of the clinical work associating meaning in life and well-being is that psychopathology may arise from lack of meaning and, conversely, attainment of meaning is healing. From this perspective, acquiring or losing meaning in life can influence positive and negative psychological functioning (Zika & Chamberlain, 1992).

Theoretically, the association between meaning and well-being derives from humanistic and existential perspectives. Humanistic theor- ists, such as Allport and Maslow, argue that people have a capacity for self-detachment and self-transcendence (Frankl, 1969). These theories equated attaining meaning with mental health, and failure to do so with mental illness. When meaning in life is not attained, a sense of meaningless is evident and an existential vacuum with associated feelings of boredom, apathy and emptiness is experienced (Frankl, 1969). Existential neurosis may arise whereby the vacuum is filled with symptoms such as alcoholism and depression (Frankl, 1987).

Existential philosophers also emphasise the importance of meaning in life and argue that meaning can be attained despite extreme or pathogenic circumstances. Frankl believed life can always have meaning, even in the Nazi concentration camps, and that the search for meaning is an inherited and dominant drive of human experience (Frankl, 1948/1997). Current literature supports Frankl's idea that meaning in life can be distilled from a sense of meaningless when it is the result of traumatic events and survivors are faced with losses (Giovinco & McDougald, 1994; Hamera & Shontz, 1978; Janoff-Bulman & McPherson Frantz, 1997; Lukas, 1998; Magomed-Eminov, 1997; Thompson, 1991). The outcome of trauma can be more than survival; value and meaning are created *because* of the profound loss, particularly of deeply held illusions such as safety (Janoff-Bulman & Berg, 1998).

Despite literature linking meaning in life to an ability to overcome negative and even highly traumatic experiences, empirical investigations into the relationship between meaning and well-being have often failed to incorporate positive and negative measures of well-being. Like meaning in life, well-being is a multidimensional construct and includes life satisfaction and positive and negative affect (Debats, 1996). Psychiatric and psychosomatic research calls into question the *DSM-IV* (American Psychiatric Association, 1994) assumption that well-being and distress are mutually exclusive (Fava, 1999). Empirical studies suggest that positive and negative components of well-being are influenced by different factors and are not just opposite ends of a continuum (Chamberlain, 1988). What makes people happy is not the same as what makes them unhappy, and vice versa (Chamberlain, 1988).

Meaning in life has been associated with a number of positive (e.g., happiness, self-esteem, spirituality) and negative (e.g., depression) well-being constructs. A positive relationship between meaning and life satisfaction and happiness has been reported in numerous studies. This relationship has been found using various measures of meaning in life (e.g., Bhogle & Prakash, 1993; Debats, van der Lubbe & Wezeman, 1993; Chamberlain & Zika, 1988a; Hutzell, 1988; Zika & Chamberlain, 1987, 1992) and in samples of students (Debats et al., 1993), the general population (Debats et al., 1993), psychiatric and non-psychotic therapeutic patients (Debats, 1996), non-working women, mothers and older women (Zika & Chamberlain, 1992).

Theorists such as Frankl, Maslow and Erickson posited a positive relationship between perceived meaning in life and self-esteem (Frankl, 1969), though high self-esteem is not sufficient to guarantee high meaning scores (Battista & Almond, 1973) and may be a function of the pursuit of goals associated with meaning (Orbach, Illuz & Rosenheim, 1987). To achieve meaning one must have high self-esteem, be internally integrated, possess a sense of control over life, and form a satisfactory relationship with the environment (Battista & Almond, 1973). A significant relationship between meaning and self-esteem has been reported in a number of studies including those of non-psychotic therapeutic patients (Debats, 1996), students (Battista & Almond, 1973; Lindeman & Verkasalo, 1996) and older adults (Reker, 1994; van Selm & Dittmann-Kohli, 1998).

A negative relationship between meaning in life and negative affect, which extends to the severe negative affect typical of depression, is a common clinical observation with depressed people frequently stating, "My life is meaningless." This observation has been reported in many empirical studies and is supported by the epidemiological findings on the link between suicide and meaninglessness. Durkheim (1897/1951) argued that loss of meaning in life could predispose to suicidal tendencies and this contention has acquired greater force given the observation of high and rising suicide rates among population subgroups whose way of life no longer readily confers meaning. Examples of these groups include young males in rural Australia (Burnley, 1995) and many Aboriginal groups.

Tribes whose languages had no words for suicide now show a very high incidence of the problem possibly due to the loss of tribal cultures (Eastwell, 1982). Studies linking meaning to depression have used various measures of meaning (e.g., Debats, 1996; Korotkov, 1998; Reker, 1994; van Selm & Dittmann-Kohli, 1998) and qualitative measures (Debats, Drost, & Hansen, 1995; Moore, 1997) in samples of students (e.g., Debats et al., 1995), the general population (Debats et al., 1993), older adults (van Selm & Dittmann-Kohli, 1998), non-psychotic therapeutic patients (Debats, 1996), and psychiatric resident in patients (Moore, 1997). It has been suggested that the strong correlation between meaning and depression found in previous studies suggests current measures of meaning in life are not independent of depression (Dyck, 1987).

The current study will examine the relative importance of positive and negative measures of well-being in predicting cognitive (Framework) and affective (Fulfillment) meaning in life. Well-being will be measured using life satisfaction (happiness), positive affect (self-esteem), negative affect (psychiatric symptoms) and the Mental, Physical, Spiritual Well-being scale (Vella-Brodrick & Allen, 1995), a multidimensional measure of well-being that contains a spiritual dimension relevant to meaning in life. The study will also examine the relationship between meaning and age and gender.

The relationship between meaning in life and gender and age is unclear. Some studies reported higher meaning scores in men compared with women (e.g., Crumbaugh, 1968; Orbach et al., 1987) and others failed to detect differences in meaning scores (e.g., Baum & Stewart, 1990; Debats, 1999; Harlow, Newcomb, & Bentler, 1986). Examination of age differences in meaning has also yielded divergent results with some studies reporting no differences in meaning scores (Debats, 1998; Reker et al., 1987) and other studies finding that meaning scores increase with age (Meier & Edwards, 1974; Van Ranst & Marcoen, 1997). Significantly higher meaning scores for middle-aged men compared with younger and older men has also been reported (Baum & Stewart, 1990). By contrast cross-sectional studies found that meaning in older adults was lower than that of middle aged and younger people (Ryff & Singer, 1998). Whether this is an age or cohort change remains unclear.

METHOD

Participants

Questionnaires were distributed to a convenience sample of 100 adults in the general population with 83 participants returning questionnaires by reply paid post (*n* for females = 57, *n* for males = 25;

one with gender not specified). Ages ranged between 18 and 84 years (*M* = 37.4, *SD* = 13.1; *M* for females = 36.7, *M* for males = 38.6).

Materials

Questionnaires presented measures in the following order: *Happiness Index*, the *Life Regard Index*, the *Mental Physical Spiritual Well-being Scale*, the *Coopersmith Self-Esteem Inventory* and the *Brief Symptom Inventory*.

The Happiness Index (Veenhoven, 1984) consists of two questions: (1) "To what extent do you regard yourself as a happy person?" and (2) "How satisfied are you with your present life?". Responses are indicated on a four-point scale (1 = Not that satisfied or Happy, 2 = Doubtful satisfied or happy, 3 = Satisfied or happy, 4 = Very satisfied or happy) which when summed comprise the total *Happiness Index*, with a high score indicating a high level of life satisfaction. Previous studies using this scale have not provided reliability coefficients. It was nevertheless selected to facilitate comparison with previous studies (Debats, 1990, 1996; Debats et al., 1995; Debats et al., 1993).

The *Life Regard Index* meaning was assessed using an English translation of the *Life Regard Index* (Debats, 1998) adapted from the original *Life Regard Index* (Battista & Almond, 1973). It contains 28 items with two subscales, Framework and Fulfillment. Each scale has 14 items, half phrased positively (e.g., "I have really come to terms with what's important for me in my life") and half negatively (e.g., "Nothing very outstanding ever seems to happen to me"). Responses to statements are indicated on a three-point scale (1 = don't agree, 2 = no opinion, 3 = agree). Subscale scores and total Life Regard Index scores are obtained by summing responses with a high score indicating a fulfilling and meaningful life.

Social desirability accounts for only four per cent of the variance of *Life Regard Index* scores (Battista & Almond, 1973). The *Index* has a reported re-test reliability of .94 (Battista & Almond, 1973) and internal consistency of .86 (Debats, 1990) and of .91 (Debats et al., 1993). Cronbach's alpha ranged between ranged between .75 and .84 for the Framework subscale and between .84 and .87 for the Fulfillment subscale (Chamberlain & Zika, 1988a; Debats, et al., 1993).

Factorial analyses of the *Life Regard Index* have been exploratory (Chamberlain & Zika, 1988b; Debats, 1990; Orbach, et al., 1987) and confirmatory (Debats et al., 1993) with most studies supporting its theoretical structure. An exception was Chamberlain and Zika (1988b) who found little empirical support for the Framework dimension in a sample of 194 women with young children. Content validity has been supported by qualitative-quantitative studies (Debats, 1998). Construct validity of the *Life Regard Index* has been demonstrated in studies reporting a positive relationship between *Index* scores and fulfillment of ultimate life goals (Battista & Almond, 1973), integration of personality (Orbach et al., 1987), effective coping with past crises (Debats et al., 1995), and emotional and other measures of well-being (Debats, 1990, 1996; Zika & Chamberlain, 1992). Finally, *Index* scores discriminated between distressed and non-distressed people and have been found to be independent of value orientations (Debats et al., 1993).

The Mental, Physical and Spiritual Well-being Scale (Vella-Brodrick, 1995) contains 30 items, 10 for each subscale. The Mental asks about mental functioning, e.g., "Do you engage in games which are designed for mental stimulation (e.g., bridge, crosswords, chess, etc.)." The Physical sub-scale focuses on aspects of health (e.g., "Over the past year have you experienced aches and pains?"). Spiritual sub-scale items contain religious and existential items such as, "Over the past year have you tried to enhance your personal or spiritual development (e.g., meditation, yoga, praying, etc.)". The existential dimensions focus on development of self, ethics, morality, self-actualisation, self-worth, life purpose, life satisfaction and fulfillment and peace within oneself (Vella-Brodrick, 1995).

Responses are indicated on a five-point scale, ranging between often and never with half the items reverse scored. Scores within each subscale are added. The subscale scores are more meaningful than the overall well-being score as they indicate strengths and weaknesses in particular areas (Vella-Brodrick, 1995).

Correlations between the Mental, Physical and Spiritual subscales and the *Marlowe-Crowne* were low (Mental = .06, Physical = -.06, Spiritual = -.04) indicating the scale does not elicit socially desirable responses (Vella-Brodrick & White, 1997). Internal consistency (Mental, $\alpha = .75$; Physical, $\alpha = .81$; Spiritual, $\alpha = .85$) and test-retest reliability (Mental, $r = .94$; Physical, $r = .87$; Spiritual, $r = .97$) taken over a one-month interval support reliability of the scales (Vella-Brodrick & Allen, 1995). Cross validation in a sample of university and employees in a research organization support the three-factor structure of the scale (Vella-Brodrick & Allen, 1995). Scores on the Mental and Physical subscales were significantly, but slightly, correlated with the General Health Questionnaires providing evidence of concurrent validity (Vella-Brodrick & Allen, 1995). The Mental and Spiritual subscales significantly discriminated three criterion groups nominated as high on mental, spiritual and physical activities (chess players, weight-trainers, prayer group participants) supporting the discriminant validity of the Mental Physical Spiritual scale (Vella-Brodrick & Allen, 1995).

The *Coopersmith Self-Esteem Inventory* (Adult Form), is designed to measure attitudes towards the self in social, academic, family and personal areas of experience (Coopersmith, 1987). Self-esteem refers to self-evaluation of worthiness, approval, competence, success and significance. The *Inventory* consists of 25 negative and positive items. Responses are indicated by ticking either "like me" or "unlike me". A point is given for each "like me" response on the positive items and for each "unlike me" on the negative items (maximum = 25). High scores indicate high self-esteem. Internal consistency for the Adult Form of the inventory was .81 (Coopersmith, 1981) with several studies supporting the construct validity of the scale (Coopersmith, 1987). The scale had a strong negative correlation with measures of anxiety and low to moderate negative correlations with social desirability providing moderate support for divergent validity of the scale (Coopersmith, 1987). The *Inventory* has been found to be significantly correlated with behavioral observations and other measures of self-esteem providing strong support for its convergent validity.

The *Brief Symptom Inventory* is a 53-item, self-report inventory designed to measure the psychological symptoms of psychiatric and medical patients as well as non-patients (Derogatis, 1993). Items are statements about psychological symptoms and respondents indicate how distressed they were within the last seven days using a five-point scale (0 = not at all, 4 = extremely).

There are nine subscales: Somatisation, Obsessive-Compulsive, Sensitivity, Depression, Anxiety, Hostility, Phobic Anxiety, Paranoid Ideation, Psychoticism. There is an additional scale (for clinical use and used in global indices) and three Global indices. The users manual reports high reliabilities for the subscales, ranging between .71 and .85 and excellent convergent and divergent validity (Derogatis, 1993). Construct validity was supported with examination of the internal structure of the scales revealing only minor differences between the empirical factor structure and the hypothesized *a priori* structure (Derogatis, 1993).

Procedure

All procedures were approved by The Standing Committee on Ethics in Research on Humans, Monash University. People who agreed to participate were asked to take away and complete the questionnaires in their own time, and return them within one week. They were informed that the task would take approximately 30 minutes and that responses would be anonymous, and were to be returned in a reply paid envelope.

RESULTS

Statistical analysis of the data was conducted using SPSS for Windows. Table 1 shows that Cronbach's Alpha reliabilities for all scales was strong ranging between .76 (Physical) and .93 (Life Regard Index). All *Life Regard Index* scales were slightly higher than those reported by previous studies. Compared with previously reported reliabilities of the *Mental Physical Spiritual Well-Being Scale* (Vella-Brodrick & Allen, 1995) reliabilities in Table 1 are lower for the Mental, higher for the Physical and similar for the Spiritual subscales. Both the Self-esteem Inventory and *Brief Symptom Inventory* scales yielded similar reliabilities when compared to those reported by previous authors (Coopersmith, 1987; Derogatis, 1993). Although there are no suitable data to compare the Cronbach's alpha for the *Happiness Scale*, the current findings suggest that it is reliable.

TABLE 1 Alpha Reliabilities of Meaning in Life and Well-Being Measures

Scale	M	SD	Alpha	No. Items	N
Life Regard Index (Overall)	66.81	11.43	.93	28	79
Framework	33.84	5.35	.85	14	83
Fulfillment	33.06	6.91	.88	14	79
Happiness Index	6.01	1.47	.82	2	81
MPS Well-Being Scale					
Mental	34.16	4.94	.51	10	82
Physical	34.23	7.09	.76	10	82
Spiritual	29.31	8.72	.86	10	82
Self-Esteem	18.03	4.74	.82	25	82
Brief Symptom Inventory					
Somatization	3.19	3.86	.80	7	81
Obsessive-Compulsive	6.17	4.22	.79	6	82
Interpersonal Sensitivity	3.92	3.89	.87	4	82
Depression	.75	.92	.91	6	82
Anxiety	4.27	3.96	.80	6	81
Hostility	3.73	4.11	.87	5	82
Phobic Anxiety	1.17	2.25	.77	5	82
Paranoid Ideation	3.54	3.61	.78	5	82
Psychoticism	2.57	3.52	.81	5	81

Student's *t* tests for independent groups revealed that there were no gender differences in Fulfillment ($t_{76} = -.99$) or Framework ($t_{80} = -.01$). To examine age differences in well-being measures, age was recoded to form three groups of approximately equal size (< 31 years, 31 – 42 years, > 42 years) and one way Analysis of Variance was performed. Where the assumption of homogeneity of variance was met, Scheffé's post hoc test was employed. Where homogeneity was violated, Dunnett's C was calculated as this test does not assume homogeneity of variance. The youngest age group had significantly lower Framework scores than either of the other two age groups. Mean scores on the Framework for the three age groups from youngest to oldest were 31.70, 35.56 and 34.15 respectively. The youngest age group was significantly lower on Framework ($F_{(2,77)} = 3.98$ $p < .05$). Mean scores on the Fulfillment for the three age groups from youngest to oldest were 32.15, 32.52 and 34.27 respectively. There were no significant differences by age on Fulfillment with $F_{(2,73)} = .71$.

Pearson product-moment correlations between all measures were obtained, though unfortunately it has not been possible to show these for reasons of space. As expected, Framework and Fulfillment scales are significantly related to each other and each scale is significantly related to all but two well-being measures (Mental, Somatization). Although three positive well-being measures (Happiness, Self-esteem, Physical) are significantly related to negative well-being subscales, the distinction between positive and negative well-being was evident in the lack of any significant relationship between two positive (Mental, Spiritual) and all the negative well-being measures.

The non-significant relationship between the Mental, Physical and Spiritual subscales indicates that the Mental, Physical and Spiritual Well-being Scale taps into three separate dimensions of positive well-being. In contrast, The Brief Symptom Inventory appears to measure a highly correlated "cluster" of negative components.

Since only a few correlations were less than .7, the criteron for possible multicollinearity (Tabachnick & Fidell, 1996), two hierarchical regression analyses were conducted to determine if positive well-being improved prediction of Framework and Fulfillment beyond that afforded by negative well-being measures. Although there were only 77 cases after removal of cases with missing data, the minimum requirement for the ratio of cases to independent variables was satisfied (Coakes & Steed, 2001). For both regression analyses no multivariate outliers among the cases were identified (criterion for Mahalanobis distance $p < .001$). Examination of residual scatterplots revealed that assumptions of normality, linearity, homoscedasticity and independence of residuals were not violated. All requested variables were entered, indicating no multicollinearity or singularity.

Results of the hierarchical regression using positive and negative well-being measures to predict Framework found *R* to be significantly different from zero at the end of each step (Table 2). After the first step, 24 per cent of the variance in Framework was explained by the negative well-being predictors with two measures (Depression, Obsessive-compulsive) being significant predictors. After the second step, examination of the R^2 Change statistic indicates that well-being measures made a significant unique contribution of 26 per cent to the explained variance of Framework, with two positive well-being measures (Happiness, Spiritual) significantly predicting Framework. However, examination of the significance of the predictors indicates that one negative well-being measure (Obsessive-compulsive) is not a significant predictor when all well-being variables are entered into the regression equation. Obsessive-compulsive is a significant predictor of Framework only when considered in combination with other negative well-being measures, but not when positive measures of well being are included.

TABLE 2 Hierarchical Multiple Regression of Positive and Negative Well-being Measures on Framework (*n* = 77)

	B	*β*	Summary Statistics	
Framework (dependent variable)				
Model 1				
Somatization	.29	.21		
Obsessive-Compulsive	- .58**	-.46	R^2	= .33
Interpersonal Sensitivity	.40	.29	Adjusted R^2	= .24
Depression	-..96***	-.98	*R*	= .57***
Anxiety	.32	.24		
Hostility	-.006	-.05	R^2 change	= .33***
Phobic Anxiety	-.25	-.11	*F* change (*df* = 9, 67)=	3.61***
Paranoid Ideation	-.10	-.07		
Psychoticism	.75	.47		
Model 2				
Somatization	.17	.12		
Obsessive-Compulsive	-.36	-.29	R^2	= .59
Interpersonal Sensitivity	.20	.14	Adjusted R^2	= .49
Depression	-.60*	-.61	*R*	= .77***
Anxiety	.24	.17		
Hostility	.18	.14		
Phobic Anxiety	-.26	-.11	R^2 change	= .26***
Paranoid Ideation	.0006	.00	*F* change (*df* = 5, 62)=	7.75***
Psychoticism	.54	.35		
Happiness	.87*	.23		
Mental	.009	.08		
Physical	.003	.04		
Spiritual	.25***	.41		
Self-esteem	.21	.18		

* $p < .05$. ** $p < .01$. *** $p < .001$

Initial analysis of the hierarchical regression using positive and negative well-being measures to predict Fulfillment revealed that Interpersonal Sensitivity was a suppressor variable. This subscale loaded on four separate dimensions in recent factor analyses of the *Brief Symptom Inventory* items (Derogatis, 1993). Derogatis (1993) argued that it seems likely that the failure of this subscale to form a coherent factor may be due either to some peculiarity of his sample composition or that the subscale may have too few items to sustain invariance across important population parameters. Interpersonal Sensitivity was, therefore, removed from the analysis. Results of the hierarchical regression using all other well-being measures to predict Fulfillment revealed *R* to be significantly different from zero at the end of each step (Table 3). After the first step, 33 per cent of the variance in Fulfillment was explained by the negative well-being predictors with two measures (Depression, Obsessive-Compulsive) being significant predictors. After the second

TABLE 3 Hierarchical Multiple Regression of Positive and Negative Well-being Measures on Fulfillment (n = 77)

	B	β	Summary Statistics	
Fulfillment (dependent variable)				
Model 1				
Somatization	.27	.15		
Obsessive-Compulsive	-.75**	-.45	R^2	= .40
Depression	-1.19***	-.92	Adjusted R^2	= .33
Anxiety	.31	.17	*R*	= .64***
Hostility	.22	.13		
Phobic Anxiety	.68	.22	R^2 change	= .40***
Paranoid Ideation	-.24	-.13	*F* change (*df* = 8, 68)	= 5.74***
Psychoticism	.75	.36		
Model 2				
Somatization	8.56E-2	.04		
Obsessive-Compulsive	-.34	-.20	R^2	= .71
Depression	-.50	-.39	Adjusted R^2	= .65
Anxiety	.24	.14	*R*	= .84***
Hostility	.32	.19		
Phobic Anxiety	.38	.12	R^2 change	= .31***
Paranoid Ideation	4.32E-02	.02	*F* change (*df* = 5, 63)	= 13.13***
Psychoticism	.20	.10		
Happiness	2.50***	.51		
Mental	1.10E-02	.01		
Physical	-6.20E-02	-.06		
Spiritual	.19**	.24		
Self-esteem	.38*	.25		

* $p < .05$, ** $p < .01$, *** $p < .001$

step, examination of the R^2 Change statistic indicates that well-being measures made a significant unique contribution of 31 per cent to the explained variance of Fulfillment, with three positive well-being measures (Happiness, Spiritual, Self-esteem) significantly predicting Framework. However, examination of the significance of the predictors indicates that the negative well-being measures were no longer significant predictors when all well-being variables were entered into the regression equation. Negative well-being on its own is a salient predictor of Fulfillment. However, in combination with positive well-being, its effect is insignificant.

Comparison of the results of the two regression analysis reveal that the positive and negative well-being measures explained more variance when predicting Fulfillment than they did when used to predict Framework (Tables 2 & 3).

DISCUSSION

Positive well-being measures improved prediction of Framework and Fulfillment beyond that afforded by negative well-being measures. Cognitive (Framework) and affective (Fulfillment) meaning in life were associated with two positive (happiness, spiritual) and one negative (depression) well-being measure. Fulfillment was associated with three positive (happiness, spiritual, self-esteem) and no negative well-being measures when both positive and negative well-being were considered together. These results support the proposition that positive and negative well-being are influenced by different factors (Chamberlain, 1988) and highlights the importance of including positive and negative measures of well-being in future research examining the relation between meaning in life and well-being.

Comparison of the two regression analyses revealed that well-being measures were more strongly associated with affective (Fulfillment) meaning than with cognitive (Framework) meaning. Perhaps a sense of meaning in life is more important than a fully developed structure for that meaning. Qualitative studies report that relationships are the most frequently cited source of meaning in persons' lives (Debats, 1999). The importance of relationships is also reflected in clinical literature. When clients report that a positive relationship exists with the counsellor, this single factor accounts for more variance, in terms of positive client outcomes, than any other factor (Miller, Duncan & Hubble, 1997). Relationships are primarily affective in nature and other important sources of meaning may also be affective. Although therapists have described the therapeutic relationship (e.g., Rogers, 1951), processes such as relationships do not have a cognitive structure that is easy to articulate. Relationships and other sources of meaning may be experienced as fulfilling, even when the person cannot provide a cognitive framework for that meaning.

The association between positive well-being and meaning in life found in the current sample of normal adults is consistent with studies examining clinical samples (Debats, 1996) and mothers and people aged 60 and older (Zika & Chamberlain, 1992). The intimate relationship between positive well-being and meaning in life supports the argument that studying the role of meaning will contribute to an understanding of the determinants of well-being (Zika & Chamberlain, 1992).

Spirituality was more strongly associated with cognitive (Framework) meaning than it was with affective (Fulfillment) meaning. The Spiritual sub-scale is made up of religious and existential dimensions (Vella-Brodrick, 1995). The religious aspect focuses on a relationship with a higher being while the existential aspect focuses on development of self, ethics, morality, self-actualisation, self worth, life purpose, life satisfaction and peace with oneself. The items are not exclusive to either religious or existential dimensions of spirituality. The Spiritual sub-scale therefore includes more items that appear to be consistent with a fulfillment-like component but less with a framework-like component. However, the results indicate the opposite - the Spiritual subscale is more closely related to life purpose and goals (Framework) than with a feeling of satisfaction with life (Fulfillment). This finding may be due to the fact that the *Mental, Physical and Spiritual Well-Being* scale is based on self-reported behavior and not feelings (Vella-Brodrick, 1995). The stronger relationship of spirituality with Framework as compared with Fulfillment is therefore consistent with Framework being a cognitive, behavioral dimension and Fulfillment being an affective dimension.

The previously reported association between meaning and physical health (e.g., Eastwell, 1982; Frankl, 1969; Ryff & Singer, 1998) or between meaning and mental well-being (e.g., Moore, 1997) was not found in this study. Again this may be a function of the operational definition of mental and physical used in the Mental, Physical and Spiritual Well-Being Scale. Failure to find this association may also partly be a function of the nature of the sample. Participants were all adults engaged in paid work, tertiary study or both; which usually requires a reasonable level of physical health. This selection factor may have obscured any relationship between physical health and meaning as part of a "well worker effect."

The association between depression and cognitive meaning (Framework) is consistent with previous studies reporting that these variables are related (e.g., Debats et al., 1995; Harlow et al., 1986;

Korotkov, 1998; Recker, 1994). This consistent finding across numerous studies has led to the suggestion that current measures of meaning are not independent of depression (Dyck, 1987). Results of this study, however, suggest that, in the case of affective meaning (Fulfillment), a sense of meaning is distinct from depression since a measure of depression was not a significant predictor when positive and negative measures of well-being were combined (Table 3). Meaning in life is a conceptually broad construct, but depression is also a broad pervasive experience, having an impact on many dimensions of a sufferer's life. Finer discriminations in measures of meaning and depression may be necessary before the inter-relationships and distinctions between these two concepts can be understood.

The Obsessive-compulsive measure of negative well-being was associated with meaning when only negative measures of well-being were considered. Factor analysis of the *Brief Symptom Inventory* supports the validity of the Obsessive-compulsive subscale, a scale designed to measure a tendency to have ego-alien, unwanted thoughts (Derogatis, 1993). Results suggest that though these obsessive thoughts may interfere somewhat with feelings that life is meaningful they are not pervasive enough to interfere with feelings of happiness.

The strong link between meaning and happiness is consistent with earlier theoretical and empirical studies (e.g., Debats, 1996). Happiness was more strongly related to Fulfillment than Framework. The first happiness item appears to measure happiness as a stable characteristic (To what extent do you regard yourself as a happy person?). For people who scored high on this item, happiness may represent a way of being rather than simply an emotional reaction. Taking responsibility for the direction of your life and determining to find what is enjoyable in every day's experiences are concomitants of this decision and probably involve cognitive as well as emotional input. The second happiness item asks about current satisfaction with life and may fluctuate in the short-term (How satisfied are you with your present life?). This item is likely to be more affective and require less cognitive appraisal than the first item. The two Happiness questions probably encompass cognitive and emotional evaluations with the balance being toward the affective, especially because the questions are asking about happiness. High scores on happiness would be associated with both Framework and Fulfillment; Framework because of the need to develop a moral framework that supports a happy lifestyle and leads to Fulfillment as a consequence. To gain a greater understanding of the strong bond between happiness and meaning in life it will be necessary to use more finely discriminating measure of happiness.

The bivariate relationship between self-esteem and meaning is consistent with theoretical (Frankl, 1969) and empirical work (Debats, 1996; van Selm & Dittmann-Kohli, 1998). However, when all well-being variables were examined in combination, self-esteem was only associated with affective (Fulfillment), not cognitive (Framework), meaning in life, an association that may be the consequence of the affective nature of self-esteem. Perhaps high self-esteem is not invariably associated with happiness. Where values-oriented thinking is directed outwards towards goals and attainments in the world, self-esteem is directed inwards towards the self. This finding highlights the importance of testing multivariate relationships in any examination of the role of meaning, since univariate analyses may be inadequate to tease out the complex pattern of associations.

Similar to Baum & Stewart (1990), Debats (1998), Debats et al. (1993) and contrary to Orbach et al. (1987) results of the current study revealed no sex differences on the meaning scales. There were, however, age differences with the younger age group having significantly lower scores on Framework. This finding is in contrast with others (e.g., Baum & Stewart, 1990; Debats, 1998; Ryff & Singer, 1998) and suggests that conscious cognitive constructs for developing a Framework for meaning in life may develop with age, at least in the age range of the current sample. Overall, the *Life Regard Index* items are abstract and there are no specific examples of what constitutes a Framework for meaning in life (e.g., I feel like I have found a really significant meaning for leading my life). When a Framework for meaning can be found in external circumstances or activities (e.g., completing a degree, employment, supporting a family), it may be relatively easy to articulate a Framework for meaning. When circumstances do not facilitate the ready adoption and articulation of goals and objectives it is left to the individual to construct these meanings. People who do not readily think abstractly or have limited life experience may be unable to articulate a framework for long-term goals under these circumstances and hence score low on the *Framework* scale. It may be necessary to include the role of abstract thinking in any examination of the relationship between age and meaning in life, especially Framework.

The current sample had higher mean scores on Framework and Fulfillment than participants in the Debats (1996) study. Apparently, meaning in life is consistently important to a wide range of people in a variety of situations. Affective meaning in life (Fulfillment) was associated with positive well-being (happiness, spiritual, self-esteem) but not with negative well-being. Fulfillment was also more strongly related to well-being than was Framework. Perhaps, in terms of well-being, meaning in life is not so much about formulating goals and having a structure for meaning, rather it is concerned with doing what feels meaningful. The dynamics of the relationship between meaning, happiness and spirituality await future theoretical and empirical research, especially longitudinal examination of these relationships. The rising social interest in alternative religious movements as a source of spiritual meaning in life emphasizes the importance of research in this area.

REFERENCES

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Antonovsky, A. (1979). *Health, stress and coping.* San Francisco: Jossey-Boss.

Antonovsky, A. (1987). *Unraveling the mystery of health: How people manage stress and stay well.* San Francisco: Jossey-Bass.

Antonovsky, A. (1996). The sense of coherence: An historical and future perspective. *Israel Journal of Medical Science, 36,* 725-733.

Battista, J., & Almond, R. (1973). The development of meaning in life. *Psychiatry, 36,* 409-427.

Baum, S. K., & Stewart, R. B. (1990). *Sources of meaning through the lifespan. Psychological Reports, 67,* 3-14.

Bhogle, S., & Prakash, I. J. (1993). Indicators of subjective well-being in a non clinical adult sample. *Psychological Studies, 38,* 135-141.

Burnley, I. (1995). Socioeconomic and spatial differentials in mortality and means of committing suicide in New South Wales, Australia, 1985 – 91. *Social Science and Medicine, 41,* 687–698.

Chamberlain, K. (1988). On the structure of subjective well-being. *Social Indicators Research, 20,* 581-604.

Chamberlain, K., & Zika, S. (1988a). Religiosity, life meaning and well-being: Some relationships in a sample of women. *Journal for the Scientific Study of Religion, 27,* 411-420.

Chamberlain, K., & Zika, S. (1988b). Measuring meaning in life: An examination of three scales. *Personality and Individual Differences, 9,* 589-596.

Coakes, S. J., & Steed, L. G. (2001). *SPSS: Analysis without anguish.* Milton, Qld, Australia: John Wiley & Sons.

Coopersmith, S. (1981). *SEI: Self-Esteem inventories.* Palo Alto, CA: Consulting Psychologists Press.

Coopersmith, S. (1987). *SEI: Self-Esteem inventories* (5th printing). Palo Alto, CA: Consulting Psychologists Press.

Crumbaugh, J. C. (1968). Cross-validation of *Purpose-in-life* test based on Frankl's concepts. *Journal of Individual Psychology, 24,* 74-81.

Debats, D. L. (1990). The *Life Regard Index*: Reliability and validity. *Psychological Reports, 67,* 27-34.

Debats, D. L. (1996). Meaning in life: Clinical relevance and predictive power. *British Journal of Clinical Psychology, 35,* 503-516.

Debats, D. L. (1998). Measurement of personal meaning: The psychometric properties of the Life Regard Index. In P. T. P. Wong & P. S. Fry (Eds.), *The* *human quest for meaning: A handbook of psychological research and clinical applications* (pp. 237-259). Mahwah, NJ: Lawrence Erlbaum.

Debats, D. L. (1999). Sources of meaning: An investigation of significant commitments in life. *Journal of Humanistic Psychology, 39*, 30-57.

Debats, D. L., Drost, J., & Hansen, P. (1995). Experiences of meaning in life: A combined qualitative and quantitative approach. *British Journal of Psychology, 86*, 359-375.

Debats, D. L., van der Lubbe, P. M., & Wezeman, F. R. A. (1993). On the psychometric properties of the Life Regard Index (LRI): A measure of meaningful life an evaluation in three independent samples based on the Dutch version. *Personality and Individual Differences, 14*, 337-345.

Derogatis, L. R. (1993). *BSI Brief Symptom Inventory*. Minneapolis: National Computer Systems.

Durkheim, E. (1951). *Suicide*. Glencoe: Free Press. (Original work published 1897)

Dyck, M. J. (1987). Assessing Logotherapeutic constructs: Conceptual and psychometric status of the Purpose in Life and Seeking of Noetic Goals Tests. *Clinical Psychology Review, 7*, 439-447.

Eastwell, H. (1982). Overview: Australian Aborigines. *Transcultural Psychiatric Research Review, 19*, 221-247.

Fava, G. A. (1999). Well-being therapy: Conceptual and technical issues. *Psychotherapy and Psychosomatics, 68*, 171-179.

Frankl, V. E. (1969). Self-transcendence as a human phenomenon. *Journal of Humanistic Psychology, 6*, 97-106.

Frankl, V. E. (1987). On the meaning of love. *The International Forum for Logotherapy, 10, Spring/Summer*, 5-8.

Frankl, V. E. (1997). *Man's search for ultimate meaning*. New York: Plenum Press. (Original work published 1948)

Giovinco, G., & McDougald, J. (1994). Logotherapy: A journey into meaning for people with AIDS. *The International Forum for Logotherapy, 17*, 76-81.

Hamera, E. K., & Shontz, F. C. (1978). Perceived positive and negative effects of life-threatening illness. *Journal of Psychosomatic Research, 22*, 419-424.

Harlow, L. L., Newcomb, M. D., & Bentler, P. M. (1986). Depression, self-derogation, substance abuse, and suicide ideation: Lack of purpose in life as a mediational factor. *Journal of Clinical Psychology, 42*, 5-21.

Hutzell, R. R. (1988). A review of the Purpose in Life test. *International Forum for Logotherapy, 11*, 89-101.

Janoff-Bulman, R., & Berg, M. (1998). Disillusionment and the creation of value: From traumatic losses to existential gains. In J. H. Harvey (Ed.), *Perspectives on loss: A sourcebook. Death, dying, and bereavement* (pp. 35-47). Philadelphia, PA: Brunner/Mazel.

Janoff-Bulman, R., & McPherson Frantz, C. (1997). The impact of trauma on meaning: From meaningless world to meaningful life. In M. Power, C. R. Brewin (Eds), *The transformation of meaning in psychological therapies: Integrating theory and practice* (pp. 91-106). Chichester, England: John Wiley & Sons.

Korotkov, D. (1998). The sense of coherence: Making sense out of chaos. In P. T. P. Wong & P. S. Fry (Eds.), *The human quest for meaning: A handbook of psychological research and clinical applications* (pp. 51-70). Mahwah, NJ: Lawrence Erlbaum.

Lindeman, M., & Verkasalo, M. (1996). Meaning in life for Finnish students. *Journal of Social Psychology, 136,* 645-649.

Lukas, E. (1998). The meaning in life and the goals in life for chronically ill people. In P. T. P. Wong & P. S. Fry (Eds.), *The human quest for meaning: A handbook of psychological research and clinical applications* (pp. 307-316). Mahwah, NJ: Lawrence Erlbaum.

Maddi, S. R. (1967). The existential neurosis. *Journal of Abnormal Psychology, 72,* 311-325.

Magomed-Eminov, M. S. (1997). Post-traumatic stress disorders as a loss of the meaning of life. In D. F. Halpern & A. E. Voiskounsky (Eds.), *States of mind: American and post-Soviet perspectives on contemporary issues in psychology* (pp. 238-250). New York: Oxford University Press.

Meier, A., & Edwards, H. (1974). Purpose-in-Life test: Age and sex differences. *Journal of Clinical Psychology, 30,* 384-386.

Miller, S. D., Duncan, B. L., & Hubble, M. A. (1997). *Escape from Babel.* New York: Norton.

Moore, S. L. (1997). A phenomenological study of meaning in life in suicidal older adults. *Archives of Psychiatric Nursing, 11,* 29-36.

Orbach, I., Illuz, A., & Rosenheim, E. (1987). Value systems and commitment to goals as a function of age, integration and personality, and fear of death. *International Journal of Behavioral Development, 10,* 225-239.

Reker, G. T. (1994). Logotheory and logotherapy: Challenges, opportunities, and some empirical findings. *International Forum for Logotherapy, 17,* 47-55.

Reker, G. T., Peacock, E. J., & Wong, P. T. P. (1987). Meaning and purpose in life and well-being: A life span perspective. *Journal of Gerontology, 42,* 44-49.

Rogers, C. R. (1951). *Client-centered therapy.* Boston: Houghton Mifflin.

Ruffin, J. E. (1984). The anxiety of meaninglessness. *Journal of Counseling and Development, 63,* 40-42.

Ryff, C. D., & Singer, B. (1998). The role of purpose in life and personal growth in positive human health. In P. T. P. Wong & P. S. Fry (Eds.), *The human quest for meaning: A handbook of psychological research and clinical applications* (pp. 213-235). Mahwah, NJ: Lawrence Erlbaum.

Shapiro, S. B. (1988). Purpose and meaning: A two factor theory of existence. *Psychological Reports, 63,* 287-293.

Shek, D. T. (1991). Meaning in life and psychological well-being: An empirical study using the Chinese version of the purpose in life questionnaire. *Journal of Genetic Psychology, 153,* 185-200.

Tabachnick, B. G., & Fidell, L. S. (1996). *Using multivariate statistics* (3^{rd} ed.). New York: HarperCollins.

Thompson, S. C. (1991). The search for meaning following a stroke. *Basic and Applied Social Psychology, 12,* 81-96.

Van Ranst, N., & Marcoen, A. (1997). Meaning in life of young elderly adults: An examination of the factoral validity and invariance of the Life Regard Index. *Personality and Individual Differences, 22,* 877-884.

van Selm, M., & Dittmann-Kohli, F. (1998). Meaninglessness in the second half of life: The development of a construct. *International Journal of Aging and Human Development, 47,* 81-104.

Veenhoven, R. (1984). *Conditions of happiness.* Dordrecht, Holland: D. Reidel Publishing.

Vella-Brodrick, D. A. (1995). *A user's guide to the Mental, Physical and Spiritual Well-being Scale.* Psychology Department, Monash University, Caulfield East, Victoria 3145.

Vella-Brodrick, D. A., & Allen, F. C. L. (1995). Development and psychometric validation of the Mental, Physical, and Spiritual Well-being scale. *Psychological Reports, 77,* 659-674.

Vella-Brodrick, D. A., & White, V. (1997). Response set of social desirability in relation to the Mental, Physical and Spiritual Well-being Scale. *Psychological Reports, 81,* 127-130.

Waisberg, J. L., & Porter, J. E. (1994). Purpose in life and outcome of treatment for alcohol dependence. *British Journal of Clinical Psychology, 33,* 49-63.

Yalom, I. D. (1980). *Existential psychotherapy.* New York: Basic Books.

Zika, S., & Chamberlain, K. (1987). Relation of hassles and personality to subjective well-being. *Journal of Personality and Social Psychology, 53,* 155-162.

Zika, S., & Chamberlain, K. (1992). On the relation between meaning in life and psychological well-being. *British Journal of Psychology, 83,* 133-145.

Note: Pearson product-moment correlations between all measures are available from the senior author upon request.

Validity of the Wonderlic Personnel Test as a Measure of Fluid or Crystallized Intelligence: Implications for Career Assessment

Nancy L. Bell
T. Darin Matthews
Kerry S. Lassiter
J. Patrick Leverett
The Citadel

The current study investigated whether the Wonderlic Personnel Test (WPT; Wonderlic, 1992), a brief measure of cognitive ability frequently used in personnel settings and in career assessment, is more closely associated with either the theoretical constructs of fluid or crystallized intelligence as assessed by the Kaufman Adult Intelligence Test (KAIT; Kaufman & Kaufman, 1993). Sixty-seven adults were administered these instruments in counterbalanced order. Results indicated that the WPT was significantly positively related to the KAIT Composite IQ, Crystallized IQ, and Fluid IQ scores and was equally associated with both fluid and crystallized intelligence. In addition, the mean WPT score was not significantly different from the mean KAIT IQ scores. It is concluded that the WPT measures general (composite) intelligence and may be used most effectively as an instrument to assess individuals for jobs requiring both fluid and crystallized abilities.

Practitioners often have a need for brief measures of cognitive ability. The Wonderlic Personnel Test (WPT; Wonderlic, 1992) is one such test that is widely used in pre-employment testing (Murphy, 1984). Two reasons for its popularity are its brevity of administration and ease in scoring. The WPT is a 50-item pencil and paper test that requires only 12 minutes administration time. Total scores are purported to measure intellectual ability and have been found to correlate with scores from more established tests of intelligence. It has been shown to have good concurrent validity by being highly correlated with the Full Scale IQ (FSIQ) of the Wechsler Adult Intelligence Scale (WAIS; Wechsler, 1955) and Wechsler Adult Intelligence Scale - Revised (WAIS-R; Wechsler, 1981) (e.g., Dodrill, 1981; Edinger, Shipley, Watkins, & Hammett, 1985). Only one study (Hawkins, Faraone, Pepple, Seidman, & Tsuang, 1990) has compared the WPT to the Verbal IQ (VIQ) and Performance IQ (PIQ) of the WAIS-R. Correlations were .86 and

---

Author info: Correspondence should be sent to: Nancy L. Bell, Ph.D., Psychology Dept., The Citadel, 171 Moultrie St., Charleston, SC 29409.

North American Journal of Psychology, 2002, Vol. 4, No. 1, 113-120.
© NAJP

.84 respectively, and the authors concluded that the WPT "taps intellectual functions widely and may reasonably be considered a test of global intelligence" (p. 200) with no reference to fluid or crystallized abilities. Indeed, one should be cautioned not to interpret the VIQ or PIQ as measures of crystallized or fluid abilities because research does not support this interpretation (Flanagan, McGrew, & Ortiz, 2000).

In this study, we attempted to determine whether the WPT is more closely identified with the theoretical constructs of fluid or crystallized intelligence as assessed by the Kaufman Adult Intelligence Test (KAIT; Kaufman & Kaufman, 1993). The KAIT was chosen as the criterion measure because of its robust factor analytic base that allows construct specificity between fluid and crystallized intelligence (Kaufman & Kaufman, 1993). Such a comparison may be used not only for theoretical reasons (i.e., what types of intelligence does the WPT tap) but might provide practical information for employers and job counselors, particularly for jobs that may rely more on fluid or crystallized types of intellectual abilities.

Generally, crystallized intelligence involves primarily verbal or language-based accumulated knowledge developed largely through educational and general life experiences. It includes language skills, vocabulary, the ability to listen and comprehend oral communication, a range of general, cultural (e.g., music and art), and scientific (e.g., biology, engineering) knowledge (Kaufman & Kaufman, 1993; McGrew & Flanagan, 1998). Fluid intelligence, on the other hand, refers to a person's adaptability and flexibility when faced with a relatively novel task that cannot be performed automatically and involves possibly both inductive and deductive reasoning. More specifically, it may include forming and recognizing concepts, identifying relations, perceiving relationships among patterns, drawing inferences, and comprehending implications (Kaufman & Kaufman, 1993; McGrew & Flanagan, 1998).

According to Kaufman and McLean (1998), jobs involving mechanical activities place a greater demand on fluid intelligence, while jobs involving writing or the arts, such as music and drama, require more crystallized knowledge. In addition, McGrew and Flanagan (1998) propose that occupations such as being a mathematician or a scientist are more fluid-based, while being an accountant, military leader, lawyer, or poet are more crystalized-based. Because some jobs place a greater reliance on one of these cognitive abilities than others, it would be helpful to know how a screening measure like the WPT is related to fluid or crystallized abilities. Indeed, the finding of a relationship between the WPT and either type of intellectual abilities may provide employers or career counselors with information that could facilitate placement decisions or help clients make better vocational decisions.

METHOD

Participants and Procedure

Participants were 67 adults (45 women and 22 men) who were enrolled in a small southeastern college. Their average age was 27 years (*SD* = 8.6; range 19 to 54) and the sample included 73.1% Caucasian, 14.9% African-American, 4.5% Hispanic, and 7.5% Asian and Pacific Islander adults. All students volunteered to participate in a larger study investigating the validity of brief intelligence tests. Students were administered the WPT and the KAIT in counterbalanced fashion to control for order effects. One-half of the sample was group-administered the WPT at the beginning of the semester and the other half was group-administered the WPT at the end of the semester. The KAIT was individually administered throughout the semester to these students. Scores for both instruments were generated as detailed in the test manuals. WPT raw scores were age-corrected as instructed in the test manual (1992) and standardized to a mean of 100 and standard deviation of 15 according to Dodrill's Wonderlic-to-WAIS IQ conversion table (Dodrill, 1981).

Instruments

Wonderlic Personnel Test (WPT; Wonderlic, 1992). The WPT is a written test of mental abilities, which can be administered and scored in about 15 minutes. It consists of 50 items on various types of subject matter, including general knowledge, similarities and differences in verbal and nonverbal materials, vocabulary, visual-spatial tasks, reasoning, abstraction, and math calculation. The raw score is obtained by adding the number of items correctly answered in 12 minutes. The test is designed to provide employers and career counselors with greater understanding of the cognitive ability of the people they interview, train, supervise, or counsel. The manual provides test score averages for people in various occupations and for those with different levels of education. According to the manual, higher scoring applicants will learn more quickly, master more complex material, and make better decisions with limited information. Lower scoring applicants will require more time to complete tasks, specific and constant instruction, and standardized job routines.

Internal consistency reliabilities of the WPT range from .88 to .94 while alternate-form reliability estimates range from .73 to .95 (Wonderlic, 1992). Test-retest reliabilities range from .82 to .94 (Dodrill, 1983; Wonderlic, 1992).

Kaufman Adolescent and Adult Intelligence Test (KAIT; Kaufman & Kaufman, 1993). The KAIT is a measure of general intelligence for persons aged 11 through 85 years of age. The test yields Crystallized, Fluid, and Composite IQ scores (*M* = 100; *SD* = 15), and follows the theoretical model of intelligence developed by Cattell and Horn (Cattell, 1963; Cattell & Horn, 1978; Horn & Cattell, 1966). The KAIT's Core Battery includes three Crystallized subtests (Definitions, Double Meanings, Auditory Comprehension) and three Fluid subtests (Rebus Learning, Logical Steps, Mystery Codes). The standardization sample consisted of 2,000 individuals, approximating 1990 U.S. Census data across four separate demographic strata. A large body of research, both contained in the KAIT manual and as demonstrated through several independent studies (Kaufman, 1993; Kaufman, Chen, & Kaufman, 1995; Kaufman, Kaufman, & McLean, 1995; McGrew, 1997; McGrew, Untiedt, & Flanagan, 1996) support the reliability and validity of this instrument.

Mean split-half reliability coefficients for the KAIT IQ scores ranged from .95 to .97 for the normative sample. Mean test-retest reliability coefficients based on 153 identified normal individuals in three age groups (11-19, 20-54, 55-85+ years) and retested after a 1-month interval ranged from .87 to .94 for the KAIT IQ scores (Kaufman & Kaufman, 1993).

Validity coefficients of the KAIT Composite IQ, when compared to the WAIS-R, with four samples of predominantly normal individuals from the ages of 16-19 years to 50-83 years (total N=343) ranged from .83 to .88 (Kaufman & Kaufman, 1993).

RESULTS AND DISCUSSION

Means, standard deviations, and Pearson correlation coefficients for all measures are presented in Table 1. All mean scores were within the average range of functioning. Dependent t-tests were conducted between the WPT mean score and the three KAIT IQ mean scores. The mean WPT standard score was not significantly different from the KAIT Composite IQ mean ($t_{(65)}$= 1.64, p>.05), the KAIT Fluid IQ mean ($t_{(65)}$= 1.67, p>.05), or the KAIT Crystallized IQ mean ($t_{(65)}$= .04, p>.05).

The WPT correlated positively and significantly with the KAIT Composite IQ scores (r=.66, p<.0001), Crystallized IQ scores (r=.62, p<.001), and Fluid IQ scores (r = .54, p<.001). Due to the fact that the standard deviation for our sample did not approach 15 points, it was concluded that there was a restriction of range and consequently correlations were corrected for restriction of range (Guilford & Fruchter, 1978). The same overall correlational pattern was revealed (see Table 1).

The correlation coefficients between the WPT and the Crystallized IQ (.62) and Fluid IQ (.54) were compared using a t-test for dependent correlations (Glass & Hopkins, 1984). Results of the t-test analysis indicated that the difference between the WPT/Crystallized IQ and the WPT/Fluid IQ correlations was not statistically significant ($t_{(64)}$ = 1.17, p > .05), suggesting that the WPT is associated with both fluid and crystallized intelligence, as defined by the KAIT. This is not surprising since the Fluid IQ and Crystallized IQ of the KAIT were found to be significantly correlated in our sample ($r = .44$, $p<.001$).

TABLE 1 Means, Standard Deviations, and Pearson Correlations

Measure		Mean	SD	r	r^a
WPT					
	Raw Composite Score	23.4	5.8		
	Converted IQ Score	106.8	11.2		
KAIT					
	Composite IQ	108.5	9.2	.66*	.82
	Crystallized IQ	108.5	8.9	.62*	.80
	Fluid IQ	106.8	11.2	.54*	.65

WPT = Wonderlic Personnel Test; KAIT = Kaufman Adolescent and Adult Intelligence Test.

Note: r^a Pearson correlations corrected for restriction of range.

*p <.0001

Results of statistical analysis suggest that the WPT is strongly associated with both fluid and crystallized abilities, making it a good choice of instruments for a practitioner seeking a quick assessment tool measuring general ability. However, because the WPT provides only one score, it does not provide a separate measure of fluid or crystallized abilities. Thus, if a practitioner in an employment setting needs to assess a potential employee for a job requiring more fluid ability, it is recommended that another measure, such as the *General Ability Measure of Adults* (GAMA; Naglieri & Bardos, 1997), which takes 25 minutes to administer and has been shown to be a good measure of fluid abilities (Maher, 2000), be considered. On the other hand, if a job requires more crystallized abilities, the practitioner may select the verbal scale of the *Shipley Institute of Living Scale* (Zachary, 1991) which takes only 10 minutes to administer and is promoted as a measure of crystallized ability. If the practitioner wishes to measure both fluid and crystallized abilities, it is recommended that he/she use a more expanded battery that has separate measures for each ability, such as the KAIT or the *Woodcock-Johnson Tests of Cognitive Ability* (WJ-C; Woodcock & Johnson, 1990).

Certain limitations of this study should be noted. This study was based upon the performance of a college student sample. Future studies should be conducted to explore the WPT's ability to measure both fluid and crystallized intelligence within a sample of individuals with less education. In addition, our sample included approximately twice as many females as males, thus future samples should include a more representative sample in regard to gender. We believe that our suggestions regarding using the WPT as a measure of general intelligence in pre-employment counseling and using the GAMA to measure fluid intelligence and the Shipley to assess crystallized abilities are sound. However, it must be noted that more empirical demonstration of such utility is needed. Future research exploring the validity of these instruments within such populations should be conducted. Additional validity studies of the WPT as a measure of fluid or crystallized abilities may consider delineating which items correlate with Fluid IQ and Crystallized IQ through conjoint factor analysis.

REFERENCES

Cattell, R.B. (1963). Theory of fluid and crystallized intelligence. *Journal of Educational Psychology, 54*, 1-22.

Cattell, R.B., & Horn, J.L. (1978). A check on the theory of fluid and crystallized intelligence with description of new subtest designs. *Journal of Educational Measurement, 15*, 139-164.

Dodrill, C.B. (1981). An economical method for the evaluation of general intelligence in adults. *Journal of Consulting and Clinical Psychology, 49*, 668-673.

Edinger, J.D., Shipley, R.H., Watkins, C.E., & Hammett, E.B. (1985). Validity of the *Wonderlic Personnel Test* as a brief IQ measure in psychiatric patients. *Journal of Consulting and Clinical Psychology, 53*, 937-939.

Flanagan, D.P., McGrew, K.S., & Ortiz, S.O. (2000). *The Wechsler Intelligence Scales and Gf-Gc Theory: A Contemporary Approach to Interpretation*. Needham Heights, MA: Allyn & Bacon.

Glass, G.V., & Hopkins, K.D. (1984). *Statistical methods in education and psychology* (2nd ed.). Boston: Allyn & Bacon.

Guilford, J.P., & Fruchter, B. (1978). *Fundamental statistics in psychology and education* (6th ed.). New York: McGraw-Hill.

Hawkins, K.A., Faraone, S.V., Pepple, J.R., Seidman, L.J., & Tsuang, M.T. (1990). WAIS-R validation of the *Wonderlic Personnel Test* as a brief intelligence measure in a psychiatric sample. *Psychological Assessment, 2*, 198-201.

Horn, J.L., & Cattell, R.B. (1966). Refinement and test of the theory of fluid and crystallized intelligence. *Journal of Educational Psychology, 57*, 253-270.

Kaufman, A.S. (1993). Joint exploratory factor analysis of the Kaufman Assessment Battery for Children and the Kaufman Adolescent and Adult Intelligence Test for 11- and 12- year olds. *Journal of Clinical Child Psychology, 22*, 355-364.

Kaufman, J. C., Chen, T.H., & Kaufman, A. S. (1995). Ethnic group, education, and gender differences on six horn abilities for adolescents and adults. *Journal of Psychoeducational Assessment, 13*, 49-65.

Kaufman, A.S., & Kaufman, N.L. (1993). *Manual for the Kaufman Adolescent and Adult Intelligence Test (KAIT)*. Circle Pines, MN: American Guidance Service.

Kaufman, A. S., Kaufman, J. C., & McLean, J. E. (1995). Factor structure of the *Kaufman Adolescent and Adult Intelligence Test* (KAIT) for whites, african americans, and hispanics. *Educational and Psychological Measurement, 55*(3), 365-376.

Kaufman, A.S., & McLean, J.E. (1998). An investigation into the relationship between interests and intelligence. *Journal of Clinical Psychology, 54*, 279-295.

Maher, C. (2000). *The General Ability Measure for Adults as a Measure of Fluid Intelligence: Comparison with the Kaufman Adolescent and Adult Intelligence Test.* Unpublished master's thesis, The Citadel, Charleston, South Carolina.

McGrew, K. S. (1997). Analysis of the major intelligence batteries according to a proposed comprehensive Gf-Gc framework. In D. P. Flanagan, J. L. Genshaft, & P. L. Harrison (eds.), *Contemporary Intellectual Assessment: Theories, Tests, and Issues* (pp. 92-104). New York: The Guilford Press.

McGrew, K. S., & Flanagan, D. P. (1998). *The Intelligence Test Desk Reference (ITDR): Gf-Gc Cross-Battery Assessment.* Needham Heights, MA: Allyn & Bacon.

McGrew, K. S., Untiedt, S. A., & Flanagan, D. P. (1996). General factor and uniqueness characteristics of the Kaufman Adolescent and Adult Intelligence Test (KAIT). *Journal of Psychoeducational Assessment, 14*, 208-219.

Murphy, K.R. (1984). The *Wonderlic Personnel Test.* In D.J. Keyser & R.D. Sweetland (Eds.), *Test Critiques*, Volume I (pp. 769-775). Kansas City, MO: Test Corporation of America.

Naglieri, J.A. & Bardos, A.N. (1997). *Manual for the General Abilities Measure for Adults.* Minneapolis, MN: National Computer Systems, Inc.

Wechsler, D. (1955). *Wechsler Adult Intelligence Scale.* New York: The Psychological Corporation.

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale-Revised.* New York: The Psychological Corporation.

Wonderlic, E.F. (1992). *Manual of the Wonderlic Personnel Test & Scholastic Level Exam II.* Wonderlic Personnel Test. Inc.

Woodcock, R.W. & Johnson, M.B. (1990). *Woodcock Johnson Psychoeducational Battery--Revised.* Allen, TX: DLM Teaching Resources.

Zachary, R.A. (1991). *Manual for the Shipley Institute of Living Scale.* Los Angeles, CA: Western Psychological Services.

Parental Acceptance-Rejection of Disabled Children in Non-Urban Pakistan

Zafar Afaq Ansari
International Islamic University, Malaysia

The purpose of this study was to investigate the attitudes of parents towards their disabled children. An Urdu version of Rohner's *Parental Acceptance-Rejection Questionnaire* was administered to 100 parents (50 fathers and 50 mothers) belonging to the lower middle class from two small towns of the northern Punjab in Pakistan. Of these, 80 parents (40 mothers and 40 fathers) had children with various types of disabilities: deaf, blind, physically disabled and mentally retarded (n = 20 each); while 20 parents (10 mothers and 10 fathers) had non-disabled children. The questionnaire was administered to the respondents orally due to their low level of literacy. Results indicated that the parents showed greater acceptance for their disabled children, as compared to non-disabled children, when the disability was manifest (when the children were deaf, blind or physically handicapped). When the disability was not manifest, as in case of mentally retarded children, there is no difference between the non-disabled and the disabled children. While gender of the child was not important, the gender of the parent emerged as a significant variable; the fathers showing greater acceptance and the mothers showing relatively greater rejection towards the disabled child.

Among the areas of developmental psychology that have been extensively studied cross-culturally, is the parental acceptance-rejection of children (Rohner, 1986, 1994, 1999). The Parental Acceptance-Rejection Theory (PART) postulates the central importance of parental acceptance (or *warmth*) as a factor in parent-child interaction. Warmth and affection from the most significant people in the child's environment—parents—is an important psychological need of the children, which, if unfulfilled, can result in a number of distortions in the personality of children.

Based on this theory, Rohner developed several instruments for measuring rejection or lack of warmth. These questionnaires were specifically designed to be of multi-cultural use from the onset. Therefore from conceptualization to wording an effort was made to have easy cross-

---

Author info: Correspondence should be sent to: Dr. Zafar Ansari, Dept. of Psychology, International Islamic University, 53100 Gombak, Malaysia. E-mail zafaq@yahoo.com

North American Journal of Psychology, 2002, Vol. 4, No. 1, 121-128.

© NAJP

cultural comparability and translability. The *Parental Acceptance-Rejection Questionnaires* have been translated into at least 25 languages and have been used in a number of cultures, showing great promise as useful research tools, not requiring much modification when used in countries as diverse as Korea, India, Pakistan, Egypt, and Sudan (Rohner, 1994). Reliability of retrospective reports of maternal acceptance-rejection was studied over a 7-year period (Cournoyer & Rohner, 1996). It was found that although the participants were 7-11 years old at the time of the first interview, there was a moderate agreement with their earlier report (r=.62).

Pakistani studies of parental attitude towards the handicapped have usually been concerned with mental retardation, and have been carried out in large metropolitan areas far removed from the small towns and rural areas that account for more than two-thirds of Pakistan. In a recent review of such research Shahzadi (2000), who carried out her investigations in Karachi, the largest city in Pakistan, found that parents feel a great deal of distress in raising mentally disabled children, because they "feel ashamed and embarassed about it." According to her, the most common reaction to such children is "the development of dislike for the child" (p. 104). An earlier review of the Pakistani situation by Miles & Miles (1993) yielded identical findings. Butt (1991) found that in Rawalpindi, which is another major city, the people who had had some contact with disabled children were more positive towards them. She also found that while people had a negative attitude towards mentally retarded children, those who belonged to the lower socio-economic level had a less negative attitude towards such children. It seems that this phenomenon is related to higher expectations regarding the disabled children in the upper socioeconomic class, as Shahzadi (2000) pointed out.

The present study aimed at understanding the attitude of parents towards their children, including disabled children of various types in the semi-urban settings of north Punjab in Pakistan. The questions asked were: Is there any difference in the attitude of parents towards the non-disabled children and the disabled children; and, Is there any difference between the attitude of parents towards children with different disabilities? Differences in attitudes were were studied in relation to gender of parents as well as gender of the children.

METHOD

Participants

This study was carried out as a part of a larger study in which trained researchers lived for an extended period in two small towns of northern

Punjab in Pakistan—Hasan Abdal and Taxila/Wah, and carried out extended interviews with parents of disabled and non-disabled children regarding their perceptions of disability. The sample consisted of men and women, all married, but not belonging to the same family. Fifty parents were interviewed from each of the two locations. The parents were selected on the basis of condition of their children: non-disabled, deaf, blind, physically handicapped and mentally retarded were represented equally (*n*= 10 each) from the two locations. All parents belonged to the lower or lower middle class as indicated by their area of residence. The sample design is shown in Table 1.

TABLE 1 Sample Design of the PARQ Study

Location		
	Hasan Abdal	50
	Taxila/Wah	50
Parent		
	Mother	50
	Father	50
Type of Disability		
	None (non-disabled)	20
	Deaf	20
	Blind	20
	Physically Handicapped	20
	Mentally Retarded	20

Instrument

The adult version of the *Parental Acceptance Rejection Questionnaire* (PARQ) (Rohner, 1986; 1997, Rohner, Saavedra & Granum, 1977) was used in this study. The questionnaire is meant to assess acceptance-rejection of the child, from the perspective of an adult caregiver (usually mother or father). The scale consists of 60 items measuring the four dimensions of Parental Acceptance and Rejection: Lack of Warmth, Aggression and Hostility, Neglect and Indifference, and Undifferentiated Rejection. However since the reliability of sub-scales was found to be quite low, only the total score was used for further analysis. Internal comsistency reliability (alpha) of the total test score (60 items) was found to be .86.

This questionnaire was earlier translated into Urdu by Haque (1981), and used in several studies in the urban areas of Sindh where people generally speak Urdu in their daily life and are more often literate (see for instance, Haque, 1987; Rafail & Haque, 1999; Sarfraz, 1991). In the northern area of Punjab people understand Urdu but their mother tongue is Punjabi or Pothwari, and many of the respondents are either illiterate, or minimally literate. For the purpose of this study the translation was revised by a group of psychologists who had some experience working in Pakistani rural areas, particularly in North Punjab. The revision mainly consisted of making the language of the questionnaire simpler and more readily understandable. The revised questionnaire was administered to a small group of illiterate people to find out if they were able to comprehend the meaning of items. Suitable modifications were made in the light of this experience.

Since many of the respondents were illiterate, or at least not very proficient in reading or in filling out questionnaires of this type, the test was administered orally. The response was recorded after each question was read to the respondents.

RESULTS AND DISCUSSION

The main results are shown in Table 2. Table 3 presents a three-way analysis of variance indicating the role of nature of disability, gender of the child and gender of the parents in influencing the attitude of parents towards the children.

TABLE 2: Parental Acceptance-Rejection and Various Conditions

Deaf	Blind	Physically Disabled	Mentally Retarded	Non-disabled
95.4	99.5	94.5	110.4	113.5
(14.4)	(19.7)	(13.7)	(13.8)	(16.5)

n=20 for each group; *F*= 6.17, *df* (4,95), *p* <.0002. The groups that are not significantly different from each other according to a Duncan Multiple-Range Test are underlined together.

Table 2 above indicates that in general parents show more warmth towards their disabled children, as compared to non-disabled children. A Duncan Multiple-Range test was conducted in order to see if there were any significant differences between children with different conditions. The results indicate that three types of disabilities (physical handicap, deafness, and blindness) are in one group with no significant differences in between. Two conditions (mental retardation and non-disability) fall in the other group. Parents are warmer in their attitude and less rejecting towards the first group as compared to the second. It seems that the major factor determining the attitude of the parents is visiblity or explainability of disturbance. When the problem can be attributed to some physical "illness" the parents feel more comfortable and accept the condition. This means better acceptance of physically handicapped, deaf, and blind children. On the other hand, children with mental handicaps are treated less warmly. However, it should be noted that there is no difference between mentally retarded and non-disabled children in terms of parental attitude.

Earlier studies have shown that members of communities that are tradition-based have a more positive attitude toward disability, while modernization leads to a more negative perception of disability (Reiter, Mar'i & Rosenberg, 1986). The children most rejected are the mentally retarded ones and those that are most accepted are the physically disabled. This is in line with findings elsewhere. For instance, Tangri and Verma (1992) found that mothers of mentally handicapped children reported greater 'social burden' as compared to the the mothers of physically handicapped children. Verdugo, Bermejo, & Fuertes (1995) found widespread maltreatment and physical neglect of intellectually

TABLE 3 Three-way Analysis of Variance of Total PARQ Scores

Source of Variation	Sum of Squares	*DF*	Mean Square	*F*
Main Effects	7650.30	6	1275.05	5.42***
Gender (Child)	62.24	1	62.24	.27
Disability Type	5602.23	4	1400.56	5.96***
Gender (Parent)	1464.57	1	1464.57	6.23**
Two-way Interactions	2368.00	9	263.11	1.12
Child X Type	1284.28	4	321.07	1.37
Child X Parent	1.23	1	1.23	.01
Type X Parent	1634.96	4	408.74	1.74
Three-way Interaction	831.95	3	277.32	1.18
ChildXTypeX Parent	831.95	3	277.32	1.18
Explained	100850.27	18	602.80	2.56**
Residual	19044.17	81	235.10	
Total	29894.44	99	301.97	

* $p < .05$; ** $p < .01$; *** $p < .001$

handicapped children. It has also been reported (see Taylor, 1998, for instance) that in helping situations, individuals with physical disability were treated more favorably than those individuals with mental disability.

Three-way analyses of variance were carried out with gender (boy or girl), type of disability condition (non-disabled, deaf, blind, physically handicapped and mentally retarded) and gender of the parent reporting (mother or father) as independent variables and scores on the Total PARQ as the dependent variable. Table 3 shows that perceived parental warmth was largely determined by the nature of disability.

While the gender of the child is not important, the gender of the parent appears to be quite important. The fathers were found to be more warm and accepting as compared to the mothers, who are more rejecting. This finding is in line with earlier findings by Haque (1987) and Sarfaraz (1991) that mothers are more rejecting as compared to fathers. This apparent rejection may be because of the perceived close biological link between the mothers and their children. There is considerable evidence to show that mothers are more often blamed for the problems of their offspring (Miles & Miles, 1993; Shahzadi, 2000). The mother shares greater responsibility in the upbringing of the children, their training and discipline, and when they find that children are not growing up as normal individuals they experience a sense of failure. On the other hand, women are not considered a source of authority in Pakistan (Haque, 1987). Therefore the mothers have to behave in a more authoritarian and aggressive manner, more so than the father, in order to discipline the child. Thus they perceive themselves as more strict, and their own behaviour as more rejecting than what the father would perceive if he were behaving in the same manner.

REFERENCES

Butt, F. R. (1991). *Attitude of the people towards the handicapped.* Unpublished M.Phil. thesis submitted to the National Institute of Psychology, Quaid-i-Azam University, Islamabad.

Cournoyer, D.E., & Rohner, R. P. (1996). Reliability of retrospective reports of perceived marental acceptance-rejection in childhood, *Psychological Reports, 78*, 147-150.

Haque, A. (1981). *The effects of perceived parental acceptance-rejection on personality organization in Pakistani children.* (Research Report). Hyderabad: Department of Psychology, University of Sind, Pakistan.

Haque, A. (1987). Social class differences in perceived maternal acceptance-rejection and personality dispositions among Pakistani children. In C. Kagitcibasi (Ed.) *Growth and progress in cross-cultural psychology*. Berwyn: Swets North America Inc.

Miles, M. & Miles, C. (1993). *Education and disability in cross-cultural perspective: Pakistan.* (Research Report). Peshawar, Pakistan: Peshawar Mission Hospital.

Rafail, E., & Haque, A. (1999). Relationship between perceived parental acceptance-rejection and juvenile deliquency scores. *Pakistan Journal of Psychological Research, 14*, 9-16.

Reiter, S., Mar'i, S., & Rosenberg, Y. (1986). Parental attitude toward the developmentally disabled among Arab communities in Israel: A cross-cultural study. *International Journal of Rehablitation Research,9,* 355-362.

Rohner, R.P. (1986). *The warmth dimension: Foundations of parental acceptance-rejection theory*. Beverly Hills CA: Sage.

Rohner, R.P. (1994). Patterns of parenting: The warmth dimension in world-wide perspective. In W. J. Lonner & R. S. Malpass (eds.) *Psychology and culture.* Boston: Allyn & Bacon, 113-120.

Rohner, R.P. (1997). Research instruments and software available from ROHNER RESEARCH, [On Line]. http://vm.uconn.edu/-rohner/instrum.html

Rohner, R.P. (1999). Acceptance and rejection. In D. Levinson, J. Ponzetti & P. Jorgensen (Eds.). *Encyclopaedia of human emotions*, vol 1. New York: Macmillan, 6-14.

Rohner, R.P.; Saavedra, J, M. & Granum, E. O. (1977). *Development and validation of the Parental Acceptance Rejection Questionnaire: Test manual.* Washington: Boys Town Centre for the Study of Youth Development (also published in JSAS Selected Documents in Psychology, February, 1978).

Sarfraz, S. (1991). *Gender differences in perceived maternal-acceptance-rejection among the upper middle and middle class high school students of Karachi, Pakistan.* Paper presented in the Eighth International Conference of Pakistan Psychological Association, Islamabad.

Shahzadi, S. (2000). *Disability and the family of mentally retarded children.* Karachi: Department of Special Education, University of Karachi.

Tangri, P. & Verma, P. (1992). A study of social burden felt by mothers of handicapped children. *Journal of Personality and Clinical Studies, 6*, 117-120.

Taylor, C. (1998). Factors affecting the behavior towards people with disabillities. *The Journal of Social Psychology, 138,* 766-771.

Verdugo, M.A.; Bermejo, B.G.; & Fuertes, J. (1995). The maltreatment of intellectually handicapped children and adolescents. *Child Abuse and Neglect, 19,* 205-215.

Note: The author is extremely grateful to Prof. Abdul Haque for his generous help and advice during the course of this research.

Sleep Positions and Personality: An Empirical Study

Michael Schredl
Sleep laboratory, Central Institute of Mental Health

The present study investigated the relationship between body positions at sleep onset and personality dimensions that are associated with emotions. The findings indicate that such relationships seem to exist, e.g., persons who sleep in the prone position are more anxious and less self-confident than the other sleepers.

The body posture at sleep onset is very stable from night to night (Domino & Bohn, 1980), although the sequence of sleep positions during the night does not show a consistent pattern (Lorrain & De Koninck, 1998). Based on the findings regarding the relationship between emotions and body postures during the waking state (e.g.,Flack, Laird & Cavallero, 1999; Wallbott, 1998), one might hypothesize a similar association between body position at sleep onset and personality dimensions which are connected with emotional aspects. Flack et al. (1999), for example, reported that a person instructed to sit in a specific posture (e.g., leaning the body forward, clenching the fists, etc.) experiences the specific emotion (anger, in this case) significantly more intensely than other emotions. In a similar way, Dunkell (1977) has studied the personality traits of his patients in relation to their preferred sleep positions. He differentiated four common sleep positions: semi-fetal (lying on the side), fetal (lying on the side, body curled up), prone (lying face down) and royal (lying on the back). Whereas the semi-fetal position reflects the normal and well-adjusted personality, persons who sleep in fetal or prone positions are characterized by Dunkell (1977) as being anxious. The royal position, on the other hand, is – according to Dunkell (1977) – associated with self-confidence.

However, Dunkell's evidence consists solely of selected clinical cases and his hypotheses were not tested using statistical methods. Domino and Bohn (1980) have shown that positions at sleep onset (14 positions taken from Dunkell, 1977) are very stable over time (89% exact agreement, retest after six months). Their results indicated that persons with the fetal

---

Author info: Correspondence should be sent to Dr. Michael Schredl, Schlaflabor, Zentralinstitut für Seelische Gesundheit, Postfach 12 21 20, 68072 Mannheim, e-mail: Schredl@as200.zi-mannheim.de

North American Journal of Psychology, 2002, Vol. 4, No. 1, 129-132.

© NAJP

position at sleep onset are less well-adjusted than the other sleepers. In this sample of 51 females, the prone as well as the royal positions have not been reported.

The aim of the present study was to test the hypotheses based on the clinical observations of Dunkell (1977). This was accomplished by measuring personality traits which are associated with emotions, e.g., anxiety, using quantitative methods and comparing the different sleep position groups statistically.

METHOD

The participants completed a self-developed sleep questionnaire and the German version of the *16-PF Personality Inventory* (Schneewind, Schröder & Cattell, 1983). One questionnaire item elicited the position at sleep onset as detailed as possible (posture of the body, arms and legs). These descriptions were classified into four groups: semi-fetal (lying on the side), fetal (lying on the side, body curled up), prone and royal (lying on the back). The factor H (Parmia, shyness vs. self-confidence) and the trait anxiety scale of the 16 PF (Sten values: 1 to 10) were included in the analysis. Differences between sleep positions were tested by one-tailed t-tests (selected group against all other participants).

The sample included 47 psychology students, whose mean age was 22.1 years ± 3.3. There were 32 women and 15 men.

RESULTS AND DISCUSSION

In the present sample, the most common body positions at sleep onset were the semi-fetal and fetal positions. The mean differences between the prone position group and the rest group were significant for self-confidence ($t = 3.6$, $p < .0005$) and trait anxiety ($t = -2.0$, $p < .025$). However, the differences between the fetal position group and the rest group did not reach significance (self-confidence: $t = .4$, n.s.; trait anxiety: $t = .7$, n.s.). Similarly, the royal position group did not differ from persons with other sleep positions (self-confidence: $t = 0.0$, n.s.; trait anxiety: $t = -0.6$, n.s.).

Similar to Dunkell (1977) and Domino and Bohn (1980), the semi-fetal sleep position was the most common posture at sleep onset in the present sample. Whereas the observations of Dunkell (1977) regarding the sleepers in fetal and royal positions and the findings of Domino and Bohn (1980) regarding the fetal position could not be confirmed by the present findings, the hypothesis concerning the prone position was supported. These persons are more anxious and less self-confident than persons with other sleep positions.

To summarize, this pilot study indicates that a relationship between sleep positions and trait aspects of emotions seems to exist. Future studies should include a longitudinal measurement of sleep positions (e.g., sleep diary, body posture monitors) to investigate the stability of the body position at sleep onset and measures of pre-sleep mood which might modulate the relationship between sleep position and personality dimensions. This area of research will complement the findings regarding the relationship between body posture and psychological states in the normal waking state (e.g., Flack et al., 1999).

TABLE 1 Sleep Positions and the Personality Dimensions "Self -Confidence (H)" and Trait Anxiety (16 PF)

Sleep position	Self-confidence (H)	Trait anxiety
Semi-fetal (N =17)	7.82 ± 1.70	4.53 ± 1.94
Fetal (N = 16)	6.69 ± 1.54	4.69 ± 1.58
Prone (N = 7)	4.71 ± 1.38	6.14 ± 1.57
Royal (N = 7)	6.86 ± 2.04	5.29 ± 1.70

REFERENCES

Domino, G., & Bohn, S. A. (1980). Hypnagogic exploration: sleep positions and personality. *Journal of Clinical Psychology, 36*, 760-762.

Dunkell, S. (1977). *Sleep positions: the night language of the body*. New York: Morrow.

Flack, W. F., Laird, J. D., & Cavallero, L. A. (1999). Separate and combined effects of facial expressions and bodily postures on emotional feelings. *European Journal of Social Psychology, 29*, 203-217.

Lorrain, D., & De Koninck, J. (1998). Sleep position and sleep stages: evidence of their independence. *Sleep, 15*, 335-340.

Schneewind, K. A., Schröder, G., & Cattell, R. B. (1983). *Der 16-PF-Persönlichkeit-Faktoren Test*. Bern: Hans Huber.

Wallbott, H. G. (1998). Bodily expression of emotion. *European Journal of Social Psychology, 28*, 879-896.

Factors Influencing Opinions about Urban Growth and Development: Measuring Resistance to Urban Development

James T. Kitchens
Kitchens Group, Maitland, FL

Larry Powell
University of Alabama at Birmingham

David Cookson
Kitchens Group, Maitland, FL

This study sought to identify those attitudinal factors that influenced public opinion about urban growth and development. The participants were 213 residents (70 men, 143 women) of Jefferson County and Shelby County, Alabama and 400 residents of northern California who responded to a telephone survey. The factor analysis on the first study identified nine items in four factors: (1) Benefits, (2) Traffic, (3) Community Needs and (4) Desire for Growth that accounted for 57 percent of the variance. Subsequent analysis of the data from the second study indicated that the (1) Benefits and (2) Traffic factors were relatively stable for both studies, but the other two factors were not.

Past research on urban growth and development has focused primarily on economic issues (Ciscel, 2001; Lim, 1999; Mayer & Somerville, 2000), environmental impact (Benfield, Raimi, & Chen, 1999; Johnson, 2001; Van Metre & Mahler, 2000), legal issues (Dowling, 2000), or governmental policies (Brueckner, 1997; Bradshaw & Muller, 1998; Daniels, 2001; Heim, 2001) surrounding urban sprawl. Proponents of urban growth argue that new residential and business construction makes cities vibrant (Ciscel, 2001), increases real income of the local citizens (Black & Henderson, 1999), and provides more tax revenue for local governments (Lim, 1999). Opponents argue that increased urban sprawl threatens the environment (Benfield, Raimi, & Chen, 1999), increases traffic problems (Van Metre & Mahler, 2000), leads to a diminished quality of life (Montaigne, 2000; Hirschhorn,

---

Author info: Correspondence should be sent to: Dr. Larry Powell, Communication Studies Dept., University of Alabama at Birmingham, 901 South 15th St., Birmingham, AL 35294 -2060.

North American Journal of Psychology, 2002, Vol. 4, No. 1, 133-142.

© NAJP

2001), and can have a psychological impact on local residents (Kitchin & Blades, 1997; Gallagher, 1999; Pilisuk, 2001).

Often overlooked in the controversy, though, is that urban development is also a public opinion issue. In the past few years, there has been an increased focus on the political ramifications of the issue (Leo & Beavis, 1998). An increasing number of governmental policy decisions are based on public opinion, resulting in an effect that Staley (2001) called "ballot-box zoning." Public officials, faced with an increasingly complex public issue, are moving toward making decisions that are viewed as politically popular. Unfortunately, as the issue has become more controversial, the complexity of the issue has also increased. Purcell (2000) noted that consensus attitudes are increasingly difficult to achieve, with pro-growth and pro-environment forces often at a political impasse on proposed new developments.

This study sought to identify those attitudinal factors that influenced public opinion about urban growth and development. While articulation of the issues is relatively easy, measuring the public opinion associated with these attitudes is more difficult. Some residents of an area may hold pro-development attitudes, while others may just as vehemently oppose such growth and fear its environmental consequences (Herzog & Miller, 1998). Further, there may be different reasons for holding a position, even among those who agree on the pro-development or anti-development position, as people balance environmental attitudes with personal preferences for locus of control (Allen & Ferrand, 1999). In such an political environment, it is important for researchers to have a better understanding of the psychological constructs that affect public opinions on this issue. Is it merely a two-construct, bi-polar issue between the pro-growth and pro-environment advocates? Or, are the psychological constructs behind the issue more complex? If so, how can attitudes on the issue be measured in a manner that synthesizes these elements?

This study sought to clarify that problem by identifying the attitudinal factors related to the topic. Using factor analysis, the attitudinal structure of the topic can be identified. With that information, a more precise means of measuring public opinion about urban growth and development can be constructed. Two sets of data were examined, an initial pilot study in Alabama and a follow-up analysis on data from California. The research question was: What is the factor structure of public opinions toward urban growth and development?

STUDY 1

Method

Participants. The participants in Study One were 213 residents (70 men, 143 women) of Jefferson County and Shelby County, Alabama, who responded to a telephone survey in April of 2001. Participants were selected using the following procedure. First, 8,000 names matched with telephone numbers were randomly selected in clusters of twenty names. Second, names were randomly selected from that list, contacted by telephone, and asked if they would participate in the survey. If they refused, a second number was called (alternately the one immediately above or below the original number selected). Participants were monitored to maintain gender balance in each region. A total of 686 phone calls were made of which 213 agreed to participate (response rate, 31%).

Instrument. The survey instrument consisted of 17 statements about growth and development. Responses to most items were recorded on a 5-interval Likert scale ("strongly agree, agree, unsure, disagree, strongly disagree"). Those which used different scales are identified below. The specific items were:

> Would you say taxes in your area are too high, about what you would expect, or low considering the services you receive?
>
> Over the last 5 years would you say the economy in your community has improved greatly, improved somewhat, unsure, declined somewhat, or declined greatly?
>
> This area of Alabama is growing too much and steps should be taken to slow growth.
>
> This area of Alabama is growing too slowly and growth needs to be encourage.
>
> If there's going to be more development in this area, it should be residential houses and subdivisions rather than new businesses and retail stores.
>
> Any additional growth and development in the area where I live will create too many traffic and overcrowding problems.
>
> The economic benefits of commercial development outweigh the traffic and overcrowding problems that it can create.
>
> I think attracting new businesses to this community is a good way for us to get tax money we need to support things like education and keep property taxes low.
>
> Having a new business or company build in this area will create a significant new source of tax money for the local community.
>
> Traffic in my community is a serious problem.
>
> We need more businesses in this area that will create new jobs with good pay and full benefits.
>
> Retail companies like Home Depot, WalMart, and Target do not provide the high quality jobs we need in this community.
>
> Having new business move into this area is important to the continued economic health of our community.
>
> All places change; if there has to be change would you prefer to see your community get bigger with more people and a greater number of businesses, get smaller with fewer people and fewer businesses, or remain the same size.

Statistical Analysis. Four statistical analyses were conducted on the resulting data. First, a principal component factor analysis with a varimax rotation was used to identify the factor structure. Second, the factor analysis was repeated with a principal axis analysis with a varimax rotation. A minimum eigenvalue of 1.00 was used as the criteria for establishing the presence of a factor in both analyses. Third, the items in each identified factor were subjected to alpha reliability analysis. Fourth, logistic regression was used to test the predictability effectiveness of the factors for identifying those respondents who were for and against growth and development.

Results

TABLE 1: Principal Component Analysis: Study 1

	Factor Scores			
	1	2	3	4
Factor One – Benefits				
1. I think attracting new businesses to this community is a good way for us to get tax money we need to support things like education and keep property taxes low	**.73**	.00	-.10	.22
2. Having a new business or company build in this area will create a significant new source of tax money for the local community	**.60**	-.13	-.11	.12
3. We need more businesses in this area that will create new jobs with good pay and full benefits	**.71**	-.19	.29	.00
4. Having new businesses move into this area is important to the continued economic health of our community	**.83**	-.19	.00	.00
Factor 2 – Traffic				
5. Any additional growth and development in the area where I live will create too many traffic and overcrowding problems	-.17	**.72**	.00	-.17
6. Traffic in my community is a serious problem	.00	**.73**	-.24	-.17
Factor 3 – Community Needs				
7. Over the last 5 years would you say the economy in your community has improved or declined	.14	.13	**.73**	-.18
8. Retail companies like Home Depot, WalMart and Target do not provide the high quality jobs we need in this community.	.20	.33	**.60**	-.21
Factor 4 – Desire for Growth				
9. This area is growing too slowly and growth needs to be encouraged	.24	.00	.22	**.80**
Eigenvalue	3.06	1.47	1.31	1.04
Reliability	.76	.58	.26	–
Variance accounted for	25%	38%	49%	57%
Percent of Correct Classification	89%	91%	92%	91%

Principal Component Analysis. The principal component factor analysis identified nine items in four factors: (1) Benefits, (2) Traffic, (3) Community Needs and (4) Desire for Growth that accounted for 57% of the variance. Factor 1 (Benefits) contained four items that accounted for 25% of the variance, had a reliability rating of .76, and correctly predicted 82% of the pro/anti-growth sentiments. Factor 2 (Traffic) contained two items that increased the total variance to 38 percent. The reliability of these items was .58. Adding them to the model increased predictability to 86%. Factor 3 (Community Needs) contained two items. This factor increased the total variance to 49%, but added nothing to the predictability of the model. The reliability of this factor was only .26. Factor 4 (Desire for Growth) contained only a single item. Including this item increased the variance accounted for to 57%. Predictability was increased by one percentage point, to 87%. Since it was a single-item factor, it was not possible to compute a reliability coefficient for this factor.

TABLE 2 Principal Axis Analysis: Study 1

	Factor Scores		
	1	2	3
Factor One – Benefits			
1. I think attracting new businesses to this community is a good way for us to get tax money we need to support things like education and keep property taxes low	**.62**	-.00	.19
2. Having a new business or company build in this area will create a significant new source of tax money for the local community	**.86**	-.19	-.00
3. We need more businesses in this area that will create new jobs with good pay and full benefits	**.67**	-.19	.00
Factor 2 – Traffic			
4. Any additional growth and development in the area where I live will create too many traffic and overcrowding problems.	-.16	**.62**	.00
5. Traffic in my community is a serious problem	.00	**.63**	-.10
Factor 3 – Desire for Growth			
6. This area is growing too slowly and growth needs to be encouraged.	.28	-.15	**.74**
Eigenvalue	3.09	1.53	1.39
Reliability	.74	.58	-
Variance accounted for	25%	38%	49%
Percent of Correct Classification	88%	90%	91%

Principal Axis Analysis. The principal axis analysis identified six items in three factors that accounted for 49% of the variance. Factor 1 (Benefits) contained three items, accounted for 25% of the variance, and had a reliability rating of .74. Factor 2 (Traffic) contained two items and increased the total variance to 38%. The reliability of these items was .58. Factor 3 contained only one item (Desire for Growth) and increased the total variance accounted for to 49%.

STUDY 2

The effectiveness and stability of the factor structures found in Study One were tested with a second study. The participants in Study Two were 400 residents of northern California. The methodology replicated that of Study One with the exception that the questionnaire was limited to the nine items identified in the previous study. A total of 1,698 people were contacted, of which 400 qualified for the survey and agreed to participate (response rate, 23.6%).

TABLE 3: Principal Component Analysis: Study 2

	Factor Scores	
	1	2
Factor One – Benefits		
1. I think attracting new businesses to this community is a good way for us to get tax money we need to support things like education and keep property taxes low	**.80**	.21
2. Having a new business or company build in this area will create a significant new source of tax money for the local community	**64**	.26
3. We need more businesses in this area that will create new jobs with good pay and full benefits	**.76**	.00
4. Having new businesses move into this area is important to the continued economic health of our community	**.85**	.01
Factor 2 – Traffic		
5. Any additional growth and development in the area where I live will create too many traffic and overcrowding problems.	.28	**.70**
6. Traffic in my community is a serious problem	.00	**.75**
7. Retail companies like Home Depot, WalMart and Target do not provide the high quality jobs we need in this community.	.17	**.64**
Eigenvalue	2.74	1.67
Reliability	.79	.54
Variance accounted for	31%	49%
Percent of Correct Classification	90%	90%

Results

Principal Component Analysis. The principal component factor analysis identified seven items in two factors – (1) Benefits, and (2) Traffic – that accounted for 49% of the variance. Factor 1 (Benefits) contained four items that accounted for 31 percent of the variance, had a reliability rating of .79, and correctly predicted 90% of the pro/anti-growth sentiments. Factor 2 (Traffic) contained three items that increased the total variance to 49%. The reliability of these items was .54. After adding them to the model, predictability remained at 90%.

TABLE 4: Principal Axis Analysis: Study 2

	Factor Scores	
	1	2
Factor One – Benefits		
1. Having a new business or company build in this area will create a significant new source of tax money for the local community	**.74**	.39
2. We need more businesses in this area that will create new jobs with good pay and full benefits	**.69**	.18
Factor 2 – Tax Benefits		
3. I think attracting new businesses to this community is a good way for us to get tax money we need to support things like education and keep property taxes low	.49	**.62**
Eigenvalue	3.18	1.25
Reliability	.74	–
Variance accounted for	35%	49%
Percent of Correct Classification	90%	90%

Principal Axis Analysis. The principal component factor analysis identified three items in two factors: (1) Benefits, and (2) Tax Benefits. All three items were in the single factor (Benefits) in the Principal Component Analysis. Factor 1 (Benefits) contained two items that accounted for 35% of the variance, had a reliability rating of .74, and correctly predicted 90% of the pro/anti-growth sentiments. Factor 2 contained only one item; it increased the total variance to 49%. After adding them to the model, the predictability remained at 90%.

GENERAL DISCUSSION

The first study identified nine items that could be used to measure public attitudes toward growth and development. Of these items, the six items that encompassed the two major factors (Benefits and Traffic) correctly predicted 86% of the pro/anti-growth attitudes among these participants. The first factor (Benefits) was particularly powerful as a measuring device; its four items have a high reliability rating (.76) and correctly predicted more than 80 percent of the responses. The traffic factor added to the effectiveness of the model, but had a lower level of reliability.

When this factor structure was tested in a different urban setting, only a few of the items continued to maintain factor stability. The lower reliability of the traffic factor could be influenced by several factors. As an urban growth issue, traffic is related to both lifestyle (inconvenience) and environmental (air pollution) issues (Van Metre & Mahler, 2000). Attitudes related to traffic may also be based on some factors that may not be related to urban growth and development. Such factors as road repairs, length of drive to work, and location of the respondent's place of employment could be factors influencing their response to these questions.

Finally, although, the traffic factor emerged in both studies, to some degree, it was not as strong or stable as the benefits factor in either study. Further, it was not particularly effective as a predictor of overall pro/anti-growth sentiment. That effect may have been due to differing attitudinal roles for the Benefits and Traffic factors. The Benefits factor was stable and highly effective as a predictor, primarily because it represented bi-polar attitudes on the issue. Individuals who would benefit from new growth generally supported more development; those whose lifestyles would be threatened by such growth tended to oppose it. Such a straight dichotomous response did not seem to occur with the traffic issue, i.e., while increased traffic problems tended to be associated with opposition to new growth and development, the opposite response did not occur. In other words, the perception that new growth would not increase traffic problems did not seem to be associated with support for new growth and development. Attitudinally, then, the potential for increased traffic problems was viewed as a reason to oppose new growth and development, but the absence of such a problem may not have been viewed as a reason for supporting it. Further research is needed to verify this possibility.

Overall, two major factors tended to emerge as consistent attitudinal elements – Benefits and Traffic. Respondents were more likely to support additional growth and development to the extent that they believed it would provide either economic or lifestyle benefits to an area; if they saw such growth as hurting their lifestyle or the economy of the area, they were likely to oppose it. The perception that new growth would increase traffic problems was the a source of anti-growth sentiment, but had little impact on pro-growth orientations.

Still, these six items should be highly efficient in measuring these attitudes. Most of them appeared in both studies, regardless of which factor analysis technique was used. What is still unknown, however, is their generalizability to other samples. Even within these two studies, some significant differences in factor structures were noted. The only factors that remained relatively stable through both studies and all four analyses were the benefits and traffic elements. The addition of the other three items identified in Study One should be retained, though. They offer the potential to measure attitudinal structures that may be more complex in other geographic areas. In this particular study, however, most of those other items were unnecessary. Benefits and Traffic remained the dominant factors.

Overall, though, this study is a first step in understanding public attitudes related to new growth and development. The items identified could be useful to both businesses and environmental groups who seek to understand public reception to urban sprawl and new construction. Future research should continue to examine the attitudinal structures that reflect public opinion on this issue.

REFERENCES

Allen, J. B., & Ferrand, J. L. (1999). Environmental locus of control, sympathy, and pro-environmental behavior. *Environment & Behavior, 31,* 338-353.

Benfield, K., Raimi, M., & Chen, D. (1999). *Once there were greenfields: How urban sprawl is undermining America's environment, economy, and social fabric.* Washington, DC: Natural Resources Defense Council.

Black, D., & Henderson, V. (1999). A theory of urban growth. *Journal of Political Economy, 2(107),* 252-284.

Bradshaw, T. K., & Muller, B. (1998). Impact of rapid urban growth on farmland conversion: Application of new regional land use policy. *Rural Sociology, 63,* 1-25.

Brueckner, J. K. (1997). Infrastructure financing and urban development: The economics of impact fees. *Journal of Public Economics, 66(3),* 383-404.

Ciscel, D. H. (2001). The economics of urban sprawl: Inefficiency as a core feature of metropolitan growth. *Journal of Economic Issues, 35,* 405-413.

Daniels, T. L. (2001). Coordinating opposite approaches to managing urban growth and curbing sprawl. *American Journal of Economics & Sociology, 60,* 229-243.

Dowling, T. T. (2000). Reflections on urban sprawl, smart growth, and the fifth amendment. *University of Pennsylvania Law Review, 148(3),* 873-887.

Gallagher, W. (1999). How places affect people. *Architectural Record, 187(2),* 74-79.

Heim, C. E. (2001). Leapfrogging, urban sprawl and growth management: Phoenix, 1950 - 2000. *Journal of Economics and Sociology, 60,* 245-283.

Herzog, T. R., & Miller, E. J. (1998). The role of mystery in perceived danger and environmental preferences. *Environment & Behavior, 30,* 29-41.

Hirschhorn, J. S. (2001). Environment, quality of life and urban growth in the new economy. *Environmental Quality Management, 10(3),* 1-8.

Johnson, M. P. (2001). Environmental impacts of urban sprawl: A survey of the literature and proposed research agenda. *Environmental & Planning A, 33,* 717-735.

Kitchin, R. M., & Blades, M. (1997). Relations between psychology and geography. *Environment & Behavior, 29,* 554-573.

Leo, C., & Beavis, M. A. (1998). Is urban sprawl back on the political agenda? *Urban Affairs Review, 34(2),* 179-211.

Lim, D. (1999). The relative efficiency of taxes in a two-period model of urban growth. *Journal of Urban Economics, 45*, 403-420.

Mayers, C. J., & Somerville, C. T. (2000). Residential construction: Using the growth model to estimate housing supply. *Journal of Urban Economics, 48,* 85-109.

Montaigne, F. (2000). There goes the neighborhood! *Audubon, 102(2),* 60-68.

Pilisuk, M. (2001). Ecological psychology, caring and the boundaries of the person. *Journal of Humanistic Psychology, 41(2),* 25-37.

Purcell, M. (2000). The decline of the political consensus for urban growth: Evidence form Los Angeles. *Journal of Urban Affairs, 22,* 85-100.

Staley, S. R. (2001). Ballot-box zoning, transaction cost, and urban growth. *Journal of the American Planning Association, 67,* 13-19.

Van Metre, P. C., & Mahler, B. J. (2000). Urban sprawl leaves its PAH signature. *Environmental Science and Technology, 34,* 4064-4070.

Eyewitness Testimony and the Jury Verdict

Jennifer N. Sigler and James V. Couch
James Madison University

This study examined the effect of a discredited eyewitness on jurors' verdicts. Participants (N = 192) were asked to assume the role of juror as they read one of three scenarios containing circumstantial evidence, circumstantial evidence plus eyewitness testimony for the prosecution, or circumstantial evidence for the prosecution plus eyewitness testimony along with a discredited eyewitness. Results indicated that jurors who read the scenario containing only circumstantial evidence voted guilty approximately fifty percent of the time. Participants who also read testimony from an eyewitness voted guilty more than not guilty. Jurors in the discredited eyewitness condition voted not guilty more than guilty. Voting more in favor of the defendant may have been jurors' way of punishing the prosecution for perceived trickery.

At the beginning of a trial, the scale of justice is supposed to be balanced. As jurors hear evidence presented by the prosecution and defense, this scale is often tilted one way or the other. That is, as the prosecution presents evidence against the defendant, "weight" is added on the balance to tip the scale to the guilty side, and as the defense presents information in favor of the defendant, the scale is tipped toward the not guilty side. The role of the juror is to take the evidence assimilated from both sides of the scale and arrive at a verdict, a process similar to that described in Lopes' meter model (1985).

In a study by Loftus (1974), the effect of eyewitness testimony on the scale of justice was examined. Participants were asked to assume they were jurors in a mock trial as they read a scenario about a robbery and double murder. There were three conditions in the experiment. The first group was given a scenario containing only circumstantial evidence. The second group was given a scenario that included an eyewitness for the prosecution as well as the circumstantial evidence given to the first group. The final group read a scenario that included the elements from group two, but also included testimony that discredited the prosecution eyewitness. Each participant was asked to render a verdict based on the scenario he/she read.

Author info: Correspondence should be sent to: Dr. James Couch, 1 Bromley Circle, Madison, WI 53714; e-mail jv_couch@hotmail.com.

North American Journal of Psychology, 2002, Vol. 4, No. 1, 143-148.
© NAJP

Results showed an 18 percent "guilty" rate among the jurors who were presented with circumstantial evidence alone (Loftus, 1974). Seventy-two percent of the participants who heard eyewitness testimony voted "guilty." Perhaps the most astonishing finding comes from the discredited eyewitness condition. Sixty-eight percent of the participants in the group where the eyewitness testimony was discredited still voted for conviction. Jurors persisted in their beliefs that the evidence weighed heavily against the defendant, in spite of discredited eyewitness testimony.

While some studies have confirmed Loftus' (1974) basic finding (Cavoukian, 1980; Saunders, 1981), other research has reported different results (Hatvany & Strack, 1980; Weinberg & Baron, 1982; Saunders, Vidmar, & Hewitt, 1983; Kennedy & Haygood, 1992). For example, Hatvany and Strack conducted a study that simulated a court trial and found conflicting results. In this study, participants were randomly assigned to one of five groups: a no witness control group, a pro-plaintiff eyewitness group, a pro-defense eyewitness group, a discredited pro-plaintiff eyewitness group, and a discredited pro-defense eyewitness group. In the no-witness control condition, 60% of the jurors voted for a guilty verdict (Hatvany & Strack, 1980). In the two conditions in which testimony from an eyewitness was presented, jurors voted in favor of the side with the eyewitness. When the defense eyewitness was discredited, jurors simply disregarded the testimony and guilty verdicts equaled 50%. In other words, the scale of justice tipped back toward equilibrium. However, when the prosecution eyewitness was discredited, jurors "overcorrected" for this faulty testimony by returning fewer guilty verdicts than jurors in the no-witness control condition.

Hatvany and Strack's (1980) study contradicted Loftus' (1974) findings by lending support for the discrediting effect, wherein jurors seem to discount eyewitness testimony that is discredited. Other researchers have manipulated numerous variables in an attempt to explore the effects of a discredited eyewitness on jury verdicts (Kennedy & Haygood, 1992; Saunders, Vidmar, & Hewitt, 1983; Weinberg & Baron, 1982). All of these studies have found support for the discrediting effect.

These subsequent attempts to examine the influence of eyewitness testimony on the outcome of a trial used experimental conditions that are very different from the conditions of Loftus' original study. The question is whether the results conflicting with Loftus' (1974) findings are due to variations in procedure, or whether the Loftus results are an anomaly. The present study returns to Loftus' original experiment and its simpler methodology to examine the effect of a discredited eyewitness upon a juror's verdict.

METHOD

Participants

One hundred ninety-two students from a mid-size, southern university participated in the study. Students volunteered to participate as one of several alternatives and activities for course credit. Participants were between 18 and 21 years of age.

Procedures

Upon arriving for the study, students were asked to seat themselves, and were read instructions explaining the procedure. Students then signed informed consent forms indicating that they were willing participants in the study and that they understood what their participation entailed. They then received a sheet of paper on which one of three scenarios was written. Participants at a given testing period all read the same scenario. There were between 20 and 25 subjects in each session.

TABLE 1 Frequency of Guilty and Not Guilty Verdicts for the Three Experimental Conditions

		Verdict		
		Guilty	Not Guilty	Total
Circumstantial Evidence				
	Count	29	35	64
	Percent	45.3	54.7	
Circumstantial Evidence + Eyewitness Testimony				
	Count	44	21	65
	Percent	67.7	32.3	
Circumstantial Evidence + Eyewitness Testimony + Discredited Eyewitness Testimony				
	Count	20	43	63
	Percent	31.8	68.3	

The first scenario consisted of circumstantial evidence and served as the control condition. The second scenario included the circumstantial evidence given in the first scenario as well as a prosecution eyewitness. The third scenario contained both of the elements in the second scenario and also included testimony discrediting the prosecution eyewitness. Jurors were instructed to read the scenario given and then indicate their verdicts by marking not guilty or guilty at the bottom of the sheet. There was no discussion between subjects. Each participant worked individually and arrived at his/her own verdict. After all participants had reached a verdict, they were debriefed and thanked for their participation. The three scenarios used in this study are included in the Appendix.

RESULTS

Overall, jurors voted for conviction approximately 49% of the time. Table 1 shows a frequency distribution of juror verdicts for the three conditions. A significant chi-square (2, N = 192) = 16.93, $p < .001$, indicates that the verdict and the scenario are related. Guilty and not guilty verdicts were approximately equal among participants who read the scenario containing only circumstantial evidence. Jurors who, in addition to being given circumstantial evidence, were also given eyewitness testimony voted "guilty" more often (68%) than not. However, in the third condition, when the eyewitness was discredited, jurors rendered more "not guilty" verdicts than "guilty" (32%) verdicts.

DISCUSSION

Results of the present study are consistent with those of Hatvany and Strack (1980) and reported studies that followed (Kennedy & Haygood, 1992; Saunders, Vidmar, & Hewitt, 1983; Weinberg & Baron, 1982), but they are inconsistent with Loftus' (1974) results. That is, when jurors are presented with eyewitness testimony, their subjective scale of justice tips to the guilty side. When this eyewitness testimony is discredited, however, the scale of justice does not merely return to equilibrium, but goes beyond that point, tipping more toward the "not guilty" side.

Perhaps jurors felt as though the prosecution was tricking them. Voting more in favor of the defense might have been their way of punishing the opposing side for introducing a "bad" witness. Future research might explore what goes on in the minds of jurors when this "over-correction," as Hatvany and Strack (1980) termed it, is happening. One way of assessing this phenomenon might be to ask jurors to elaborate on how they arrived at a verdict. That way, researchers could learn more about each juror's subjective scale of justice, and how and why both sides of the balance tip back and forth during a trial.

REFERENCES

Cavoukian, A. (1980). *The influence of eyewitness identification evidence.* Unpublished doctoral dissertation, University of Toronto.

Hatvany, N. & Strack, F. (1980). The impact of a discredited key witness. *Journal of Applied Social Psychology, 10*, 490-509.

Kennedy, T. D. & Haygood, R. C. (1992). The discrediting effect in eyewitness testimony. *Journal of Applied Social Psychology, 22*, 70-82.

Loftus, E. F. (1974). The incredible eyewitness. *Psychology Today, 8*, 116-119.

Lopes, L. L. (1993). Two conceptions of the juror. In R. Hastie (Ed.), *Inside the juror* (pp. 255-262). New York: Cambridge University Press.

Saunders, D. M. (1981). *Discredited eyewitness evidence and jury deliberations.* Unpublished master's thesis, University of Western Ontario.

Saunders, D. M., Vidmar, N., & Hewitt, E. C. (1983). Eyewitness testimony and the discrediting effect. In S. M. A. Lloyd-Bostock and B. R. Clifford (Eds.) *Evaluating witness evidence* (pp. 57-78). New York: Wiley.

Weinberg, H. I. & Baron, R. S. (1982). The discredible eyewitness. *Personality and Social Psychology Bulletin, 8,* 60-67.

APPENDIX

Scenario used for Circumstantial Evidence Alone Condition

On Friday, November 12, 1997, Mr. Smith, the owner of a small grocery store, was confronted by a man who demanded money from the cash register. Mr. Smith immediately handled $100 to the robber, who took the money and started walking away. Suddenly, and for no apparent reason, the robber turned and fired two shots at Mr. Smith and his five-year-old granddaughter, who was standing behind the counter. Mr. Smith and his granddaughter were hit and died instantly. Two-and-a-half hours later the police arrested a suspect who was charged with robbery and murder. A trial date was set for February 3, 1998.

The prosecution presented evidence that indicated the robber was seen running from the store and into an apartment house nearby. The defendant was a resident of this apartment house. After executing a search warrant, the police found $123 in the defendant's room. Traces of ammonia were found on the defendant's shoes. Evidence was introduced that ammonia had been used to clean the floor of the store. Paraffin tests, used to determine whether a person has gunpowder particles on his hands from firing a gun, indicated that there was a slight possibility that the defendant had fired a gun during the day.

The defendant took the witness stand in his own behalf and testified that he did not commit the crime. He testified that the money found in his room represented his savings from a two-month period. He testified that the ammonia tracing could have been obtained at a different place since he worked as a delivery man and made deliveries to a number of grocery stores in the neighborhood. He testified that he had never fired gun in his life.

Based upon this evidence and assuming that guilt must be proven "beyond a reasonable doubt" would you find the defendant....

Scenario Used for Circumstantial Evidence + Eyewitness Testimony

On Friday, November 12, 1997, Mr. Smith, the owner of a small grocery store, was confronted by a man who demanded money from the cash register. Mr. Smith immediately handled $100 to the robber, who took the money and started walking away. Suddenly, and for no apparent reason, the robber turned and fired two shots at Mr. Smith and his five-year-old granddaughter, who was standing behind the counter. Mr. Smith and his granddaughter were hit and died instantly. Two-and-a-half hours later the police arrested a suspect who was charged with robbery and murder. A trial date was set for February 3, 1998.

The prosecution presented evidence that indicated the robber was seen running from the store and into an apartment house nearby. The defendant was a resident of this apartment house. After executing a search warrant, the police found $123 in the defendant's room. Traces of ammonia were found on the defendant's shoes. Evidence was introduced that ammonia had been used to clean the floor of the store. Paraffin tests, used to determine whether a person has gunpowder particles on his hands from firing a gun, indicated that there was a slight possibility that the defendant had fired a gun during the day. A store clerk who was in the store during the robbery testified that she saw the defendant fire the shots which killed the two victims.

The defendant took the witness stand in his own behalf and testified that he did not commit the crime. He testified that the money found in his room represented his savings from a two-month period. He testified that the ammonia tracing could have been obtained at a different place since he worked as a delivery man and made deliveries to a number of grocery stores in the neighborhood. He testified that he had never fired a gun in his life.

Based upon this evidence and assuming that guilt must be proven "beyond a reasonable doubt" would you find the defendant....

Scenario Used for Circumstantial Evidence + Eyewitness Testimony + Discredited Eyewitness

On Friday, November 12, 1997, Mr. Smith, the owner of a small grocery store, was confronted by a man who demanded money from the cash register. Mr. Smith immediately handled $100 to the robber, who took the money and started walking away. Suddenly, and for no apparent reason, the robber turned and fired two shots at Mr. Smith and his five-year-old granddaughter, who was standing behind the counter. Mr. Smith and his granddaughter were hit and died instantly. Two-and-a-half hours later the police arrested a suspect who was charged with robbery and murder. A trial date was set for February 3, 1998.

The prosecution presented evidence that indicated the robber was seen running from the store and into an apartment house nearby. The defendant was a resident of this apartment house. After executing a search warrant, the police found $123 in the defendant's room. Traces of ammonia were found on the defendant's shoes. Evidence was introduced that ammonia had been used to clean the floor of the store. Paraffin tests, used to determine whether a person has gunpowder particles on his hands from firing a gun, indicated that there was a slight possibility that the defendant had fired a gun during the day. A store clerk who was in the store during the robbery testified that she saw the defendant fire the shots which killed the two victims.

The defendant took the witness stand in his own behalf and testified that he did not commit the crime. He testified that the money found in his room represented his savings from a two-month period. He testified that the ammonia tracing could have been obtained at a different place since he worked as a delivery man and made deliveries to a number of grocery stores in the neighborhood. He testified that he had never fired gun in his life. The defense showed that the store clerk had not been wearing her glasses on the day of the robbery, and since she had vision poorer than 20/400 without her glasses, she could not have possibly seen the face of the robber from where she stood.

Based upon this evidence and assuming that guilt must be proven "beyond a reasonable doubt" would you find the defendant....

Child Sexual Abuse: A Review of Definitions, Instrumentation, and Symptomology

Karen Mannon
Gloria Leitschuh
Eastern Illinois University

Issues related to the conceptualization and treatment of child sexual abuse (or sexual abuse), are reviewed by the authors. The variety of child sexual abuse/sexual abuse definitions in the literature, the properties of instruments used in the research, and reported clinical outcomes and symptoms experienced by victims are reported. Symptoms are then grouped according to the Diagnostic and Statistical Manual of Mental Disorders IV-TR (American Psychiatric Association, 2000).

The literature pertaining to child sexual abuse (CSA) and sexual abuse (SA) addresses various aspects of this form of victimization. The variety of CSA/SA definitions in the literature, the properties of instruments used in the research, and reported clinical outcomes and symptoms experienced by victims are reported. Symptoms are then grouped according to the Diagnostic and Statistical Manual of Mental Disorders-IV-TR (American Psychiatric Association, 2000).

Various types of research designs and methods comprise the CSA/SA literature, such as quantitative and retrospective designs, interviews and analyses of self-reports and parent/caretaker reports. The most common type of study reviewed in the present paper involves quantitative and retrospective designs using young adult/adult female convenience samples (Alexander & Lupfer, 1987; Briere & Runtz, 1988; Chu & Dill, 1990; Leitenberg, Greenwald, & Cado, 1992; Perrott, Morris, Martin & Romans, 1998; Peters & Range, 1995; Schaff & McCanne, 1998; Wheeler & Walton, 1987; Wolfe, Gentile & Wolfe, 1989; Wozencraft, Wagner & Pellegrin, 1991). In these studies, the data were collected via interview questionnaires and/or self-report instruments. While qualitative studies of CSA/SA are less numerous, one was included in this review. Gill and Tutty (1998) employed standard qualitative techniques such as reflection and probing. Other notable studies have been completed by using parent/caretaker reports of victim behavior (Ackerman, Dykman,

---

Author info: Correspondence should be sent to: Dr. Gloria Leitschuh, Eastern Illinois University, 600 Lincoln Ave., Charleston, IL 61920. Cfgl@eiu.edu.

North American Journal of Psychology, 2002, Vol. 4, No. 1, 149-160.
© NAJP

Jones, McPherson, & Newton, 1998; Dubowitz, Black, Harrington, & Vershoore, 1993; Freidrich, Bielke, & Urquiza, 1986; Kizer, Heston, Millsap & Pruitt, 1991; Steiger & Zanko, 1990; Wells, Adams, Ensign, McCann & Voris, 1995; White, Halpin, Santilli & Strom, 1988). The literature reviewed had commonalities in the description of the instrument(s) used and identification of clinical or behavioral outcomes.

DEFINITIONS

Inconsistent definitions of CSA/SA undoubtedly influence both the prevalence rates and the psychological disturbances reported by victims (Briere, 1992). The terms CSA and SA or both are used in the literature. The choice of terminology appears to be based on author preference. Complicating the task of identifying a common definition of CSA are variables in victim responses such as severity of abuse, availability of social support, and attributional styles regarding the causes of negative life events (Ackman, 1991; Russell, 1983, 1984; Wolfe, Gentile & Wolfe, 1989). While Beitchman, Zucker, Hood, DaCosta & Ackman's review (1992) suggested that other salient variables included age at onset, sex of child, relationship to offender, frequency and duration of abuse, type of act and use of force as parameters of abuse, it is currently not clear how these factors affect research outcomes.

Sex Abuse Definitions

The Diagnostic & Statistical Manual of Mental Disorders-IV-TR (DSM-IV-TR, 2000) does not refer to SA and correspondingly lacks substantive information. Specific references to SA in the DSM-IV-TR are reflected in V Codes, which are other conditions that may be a focus of clinical attention or Axis IV-Psychosocial and Environmental factors with no criteria sets for sexual abuse of adults or children provided.

Russell (1984) defines SA as any unwanted sexual experience before age 14, or attempted or completed rape by age 17, or any attempted or completed sexual contact between relatives before the victim turned 18. This definition would seem more appropriate as a description of CSA, due to the inclusion of the age of the victim.

Finkelhor (1979) describes SA as any and all sexual activity or contact with a child where consent is not or cannot be given, including the use of force. Finkelhor (1979) also includes deception and the level of understanding the child has of sexual activity as contributing factors.

In a more recent review of methodological problems in the research, Golfman and Padayachi (2000) suggest four major factors to be considered before an experience can be regarded as sexual abuse. These include definition of sexual contact, inclusion of noncontact sexual events and sexual invitation, inclusion of wanted and unwanted sexual experiences, and age difference between victim and perpetrator.

Child Sexual Abuse Definitions

There are no criteria in the DSM-IV-TR to characterize CSA. The criteria for Pedophilia state "The person is at least 16 years of age and at least five years older than the child or children" (APA, 2000, p. 572). The criteria set given in the DSM-IV-TR applies to the perpetrator, not the victim. Further exploration of the DSM-IV-TR yields information from V Codes. For example, Code V61.21, Sexual Abuse of a child suggests, "this category should be used when the focus of clinical attention is sexual abuse of a child" (DSM-IV-TR, 2000, p 738). Codes are simply used to identify whether the victim is a child or adult.

Friedrich, Urquiza, and Bielke (1986) defined CSA as "sexual contact with an adult, whether by force or consent, to include direct contact (intercourse) and observed contact (adult exposing self to child)" (p.50). The authors purposely excluded the age component of the Finkelhor (1979) and Russell (1983) definition, because they believed that the results would be affected. In a study examining the association between eating disorders and CSA victims, Steiger and Zanko (1990) define CSA by descriptive variables such as age, perpetrator and frequency. Rimsa et al., (1988) identify and measure emotional and somatic reactions to CSA but provide no definition of CSA. They describe results in terms of age at onset of abuse, relationship of assailant (stranger or incest) and type of assault.

Schaaf and McCann (1998) and Leitenberg et al. (1992) use a modified, yet more descriptive and explicit version of Finkelhor's (1979) SA characteristic list to describe CSA. CSA is viewed as any form of sexual contact (from touching to anal intercourse before age 15 with someone at least five years older), thus adding age parameters for both victim and perpetrator. Dubowitz, et al. (1993) have a slightly different version, defining CSA "as any inappropriate sexual contact ranging from fondling to intercourse." A variation of the age descriptor was used by Peters & Range (1995), in a study assessing suicidality in college women and men. They defined CSA as unwanted sexual experiences before the age of 12 with someone who was at least 5 years older.

In a meta-analytic study of child sexual abuse correlates, Rind, Tromovitch, & Bauserman, (1998) define CSA as "a sexual interaction involving either physical contact or no contact e.g. (exhibitionism) between either a child or adolescent and someone significantly older, or between two peers who are children or adolescents when coercion is used" (p.23). Several studies did not include any definition of CSA

(Ackerman, et al., 1998; Beitchman, et al., 1992; Chu & Dill, 1990; Gill & Tutty, 1997; Rimsza & Berg, 1988; Wells, et al., 1995; White, et al., 1988; & Wozencraft, et al., 1991).

Implications

The literature shows that researchers during the past twenty years have exercised a great deal of latitude in choosing a definition of CSA. Briere (1992) suggests that findings regarding abuse correlates must be evaluated in terms of specific definitions used. The definitions of CSA used in research design and subsequent published literature has direct and substantial impact on resulting statistical reports as well as outcomes or symptomology reported (Beichtman, et al., 1992; Briere, 1992; Golfman & Padayachi, 2000). Information from this review suggests there is no one standard or consistent definition of CSA or SA. Confusion over definition is complicated by the seemingly interchangeable use of CSA and SA by those in the field, sometimes using both terms in the same published work. Commonalities do exist in components of definitions, such as age at onset, age difference between victim and perpetrator, frequency and duration of abuse. Peters (1988) states that researchers restricting themselves to earlier definitions of CSA may report more extreme outcomes.

INSTRUMENTATION

Types

The second goal of this review is to examine the psychometric properties of instruments used in past CSA research. The purpose of this review is to highlight how the use of different instruments impacts reported outcomes. According to Briere (1992), there are three main concerns related to measurement devices in CSA research: reliability, validity and sensitivity. Interpretation of data and results of CSA studies can be difficult and/or questionable due to the use of homespun measures of unknown reliability and validity (Briere, 1988). Another issue is the lack of sensitivity of author-devised instruments to CSA specific symptoms. In addition, many of these author-devised instruments are not psychometrically sound (Beitchman, et al., 1992; Briere, 1992). This hinders identification of relationships between abuse and symptomology. The absence of proven, recognized standards of measurements related to these instruments must necessarily limit interpretation of results.

A related concern is the use of general measures of psychological function applied to victims of CSA (Briere, 1992). There are few standardized instruments devised specifically for use with CSA victims contributing to the use of generic, non-standardized and study specific instruments. CSA specific, abuse-relevant measures would assist in identifying victims, and in the development of a definition and/or criteria set specific to CSA.

Standardization

A review of the literature produced forty-one instruments (Table 1), seventeen standardized and 24 non-standardized (*Tests in Print IV*, 1994). Nine of the 24 non-standardized instruments were devised by authors specifically for the studies under consideration. Reliability, validity and psychometric soundness cannot be evaluated for the non-standardized instruments and hence caution is needed when interpreting the conclusions generated by these studies.

The significance of using a standardized instrument should be tempered by the realization that reliability, validity and psychometric soundness do not automatically equate with appropriateness for CSA symptomology. Standardized instruments, if not used in accordance with instructions and intended purpose, yield compromised or skewed results (Briere, 1992).

These findings support Briere's (1992) and Golfman and Padayachi's (2000) cautions regarding measurement issues. While this is a cursory examination of instruments using a small sample, the diversity of instrumentation used in CSA research designs and reporting of results is evident.

CLINICAL OUTCOMES

The symptoms and/or outcomes identified in the literature were compared with DSM-IV-TR criteria and are reported in Table 2. The most common outcomes reported in the CSA literature were behavioral, emotional, cognitive or physical symptoms ranging from mild to severe. It should be noted that aggression is a widely reported outcome for male CSA victims (Briere & Elliott, 1994; White et al., 1988), but rarely mentioned for female CSA victims. Few male victims were included in the sample populations and no specific assertions concerning male victims were noted. It was noted that in general males appear to react to CSA with aggressive behaviors and females with depressive behavior (Dubowitz et al., 1993; Peters & Range, 1995).

The controversial meta-analytic study of child sexual abuse correlates by Rind, et.al. (1998) indicates that "CSA does not cause intense harm on a pervasive basis regardless of gender in the college population" (p.46) and that "CSA has no inbuilt or inevitable outcome or set of emotional reactions" (p. 46). Therefore no results for the study are provided in Table 2.

TABLE 1 Instruments Used in Reviewed Articles

Standardized	Authors Reporting Research Results
Brief Symptom Checklist	Leitenberg et al. (1992)
Child Behavior Checklist	Beitchman et al. (1991) Dubowitz et al. (1993) Freidrich et al. (1986) Kiser et al. (1991) Wolfe et al. (1989)
Minnesota Child Depression Inventory	White et al. (1988) Wolfe et al. (1989) Wozencraft et al. (1991)
Dissociative Experiences Scale	Chu & Dill (1990)
Family Adaptability & Cohesion Scales III	Kiser et al. (1991)
Life Experiences Questionnaire	Chu & Dill (1990)
Millon Clinical Multi-axial Inventory	Wheeler & Walton (1987)
Minnesota Multiphasic Personality Inventory	Wheeler & Walton (1987)
Personal Inventory for Children	Kiser et al. (1991)
Present State Exam	Freidrich et al. (1986)
Rorschach	Wheeler & Walton (1982)
SCL-90-R Slosson Intelligence Test	Chu & Dill (1990)
State Trait Anxiety Scale	Beitchman et al. (1991)
Tennessee Self-Concept Scale	Wolfe et al. (1989)
Traditional Family Ideology Scale	Alexander & Lupfer (1987)

Non-Standardized Instruments	Authors Reporting Research Results
Adult Physical Trauma Questionnaire	Schaaf & McCanne (1998)
Adult Sexual Experiences Questionnaire	Schaaf & McCanne (1998)
Child Manifest Anxiety Scale	Wolfe et al. (1998)
Child Physical Trauma Scale	Schaaf & McCanne (1998)
Child Sexual Abuse Questionnaire	Peters & Range (1995)
Child Sexual Experiences Questionnaire	Schaaf & McCanne (1991)
Children's Attribution Style Questionnaire	Wofle et al. (1989)
Children's Impact of Traumatic Events	Wolfe et al. (1989)
Defense Style Questionnaire	Steiger & Gascho (1990)
Eating Attitudes Test	Steiger & Zanko (1990)
Family Environment Scale	Steiger & Zanko (1990)
Family Experiences Scale	Briere & Runtz (1988)
Fear Survey Schedule for Children	Wolfe et al. (1989)
Gambill-Rickey Assertion Inventory	Wolfe et al. (1989)
History of Victimization Form	Wolfe et al. (1989)
Hopkins Symptom Checklist	Briere & Runtz (1988)
Leary Interpersonal Checklist	Wheeler & Walton (1987)
Parent Attributes Inventory	Wheeler & Walton (1987)
PTSD Questionnaire	Schaaf & McCanne (1998)
Reasons for Living Inventory	Peters & Range (1995)

Table 1 (Continued)
Instruments Used in Reviewed Articles

Non-Standardized	Authors Reporting Research Results
Sexual Abuse Fear Subscale	Wolfe et al. (1989)
Structured Interview for Sex Abuse	Wells & McCann (1995)
Suicide Behavior Questionnaire	Peters & Range (1995)
Trauma Symptoms Checklist	Briere & Runtz (1988)

Based on this review of the literature, numerous symptoms are reported as a result of CSA/SA. This may be helpful to those who are providing CSA treatment and are required to provide diagnoses based on similar symptoms. What remains questionable is how prevalence and incidence rates are affected by the range and specificity of definition and measurement methodology.

CONCLUSIONS

The symptoms identified in the literature are not surprising given the wide range of experiences that constitute sexual abuse (Saywitz, Mannarino, Berliner, & Cohen, 2000). No single symptom characterizes the majority of those who are sexually abused (Saywitz et al.) and it is estimated that 55% of children referred for treatment have more than one diagnosis (Target & Fonagy, 1996). The effects of the abuse are also influenced by variables such as level of pre-abuse functioning and social, emotional, and financial resources (Saywitz, et al., 2000).

Clinicians and other professionals are not yet well informed regarding research and treatment modalities for sex abuse victims (Briere, 1996). Long-term outcomes for CSA victims are not well known (in part due to the lack of longitudinal studies with standardized instruments). The infamous meta-analysis by Rind, Tromovitch, and Bauserman (1998) which disputed the correlation as well as the severity of symptoms with CSA, was completed entirely with college student populations. In fact, a high rate of homelessness occurs for women who have been sexually abused. Goodman, Dutton, and Harris (1995) found that 58% of a sample of mentally ill and previously homeless women had been sexually abused as children and that 65% were sexually abused as adults. These victims were most likely not included in the college population and would have been left out of the Rind study. Future meta-analyses should include samples more representative of the CSA/SA population.

Haugaard (2000) has recently proposed three possible strategies for reducing problems created by ambiguous definitions of CSA. He suggested first that creating a narrower definition could improve the accuracy of CSA estimates. Second, a CSA definition could vary across contexts. For example, a narrower definition could be used by researchers and a broader definition could be used by clinicians. A third possibility would be to maintain the broad definition and create separate subgroups. For example CSA could be based on severity of abuse.

Haugaard (2000) acknowledges limitations to these approaches. Behavior that is abusive may depend on individual and cultural values. Definitions which change depending on context might erroneously exclude some abusers from legal prosecution. While no strategy for creating subgroups has been determined, several possibilities exist. Subgroups could be formed on the basis of the sexual act, the child's reaction to the experience, or the age of the child at the time of the sexual experience.

Recommendations for further research and study are to identify a standard definition for CSA and to develop a criteria set or symptoms-list that could be used in the DSM "V" codes. This would result in substantial improvement in treatment techniques, as well as research methodology. Development of a standardized instrument for use with the CSA population is also recommended. Until this is done, the psychometric properties of any study-specific instruments should be clearly specified in published reports.

In the absence of a clear set of sexual abuse criteria, it is suggested that clinicians provide treatment according to individual symptoms being reported and use the above findings as a guide. Clinicians will need to provide the most efficacious treatments available for screening, assessment, and treatment planning, since no single intervention will be effective for all sexual abuse victims (Saywitz et al. 2000).

TABLE 2 Symptoms/Outcomes Associated with DSM-IV-TR Diagnosis/with Authors Reporting

Axis	DSM-IV Diagnosis	Symptom/Outcome
Axis I	Depression	Low self worth, feelings of hopelessness, isolation, insomnia or hypersomnia, poor or increased appetite, increased thoughts of or attempts of suicide (Ackerman, et al., 1992; Beitchmnan, et al. ,1992; Briere & Elliott, 1994); Briere & Runtz , 1988; Dubowitz, et al., 1993; Perrott et al., 1998; Peters & Range, 1995).
	Anxiety	Derealization, distortion of reality, Fear of losing control, of being crazy, nightmares, inability to settle down, feeling on "edge," (Ackerman et al., 1998; Beitchman, et al.1992; Briere & Elliott, 1992; Briere & Runtz, 1988; Peters & Range, 1995).
	Dissociative	Feelings of detachment from one's body, distortion of reality, depersonalization, denial or repression. (Briere & Runtz ,1988; Chu & Dill, 1990; Gill & Tutty, 1997; Leitenberg et al., 1992; Perrott et al., 1998).
	Posttraumatic Stress Disorder	Distressing thoughts, images, perceptions, nightmares, restricted range of feelings, avoidance of anything related to the trauma, hypervigilance, concentration difficulties, irritability or sudden outbursts of anger, lack of interest in normally enjoyable activities (Briere & Elliott, 1994; Finkelhor, et al., 1990; Kiser, et al., 1991; Schaaf & McCanne, 1998; Wells, et al., 1995; Wolfe, et al. , 1989).
Axis III	Somatic Complaints	Abdominal pain, nausea, headaches, generalized body malaise or feeling bad, increased incidence of vaginal irritations, increase in sexually transmitted diseases & infections, gastrointestinal disorders (Ackerman, et al., 1998; Dubowitz, et al., 1993; Rimza, et al., 1988; White, et al., 1998).
Axis IV	Interpersonal Difficulties	Relationship issues (parent, child, partner, friends). Self-esteem/self-worth, control of one's life, sexual difficulties, distrust of self/others, occupational problems, school/academic problems. Lack of or poor coping skills, absenteeism/truancy or runaway behaviors (Ackerman, et al. (1998); Leitenberg, et al. (1992); Peters & Range (1995); Schaaf & McCanne (1998); White, et al. (1988).

REFERENCES

Ackerman, T., Newton, E., McPherson, W., Jones, J., & Dykman, R. (1998). Prevalence of posttraumatic stress disorder and other psychiatric diagnoses in three groups of abused children (sexual, physical, and both). *Child Abuse & Neglect, 22*(8), 759-774.

Ackman, D. (1991). A review of the short-term effects of child sexual abuse. *Child Abuse & Neglect, 15*, 537-556.

Alexander, P. & Lupfer, S. (1987). Family characteristics and long term consequences associated with sexual abuse. *Archives of Sexual Behavior,16*(3), 235-245.

American Psychiatric Association. (2000). *Diagnostic and Statistical Manual of Mental Disorders*. (4th ed.). Washington, DC: Author.

Beitchman, J., Zuker, K., Hood, J., DaCosta, G., Ackman, D. & Casavia, E. (1992). A review of the long term effects of child sexual abuse. *Child Abuse & Neglect, 16*, 101-118.

Briere, J. (1992). Methodological issues in the study of sexual abuse effects. *Journal of Consulting and Clinical Psychology, 60*(2), 196-203.

Briere, J. (Ed.). (1996). *The APSAC Handbook on Child Maltreatment.*

Briere, J. & Runtz, M. (1988). Symptomology associated with child sexual victimization in a nonclinical adult sample. *Child Abuse & Neglect, 12*, 51-59.

Briere, J. N. & Elliott, Diana M. (1994). Immediate and long-term impacts of child sexual abuse. *Sexual Abuse of Children, 4*(2), 54-69.

Chu, J. & Dill, D. (1990). Dissociative symptoms in relation to child physical and sexual abuse. *American Journal of Psychiatry*, February 1990, *147*, 887-892.

Dubowitz, H., Black, M., Harrington, D. & Vershoore, A. (1993). A follow-up study of behavioral problems associated with child sexual abuse. *Child Abuse & Neglect. 17*, 743-754.

Finkelhor, D., Hotaling, G., Lewis, I. & Smith, C. (1990). Sexual abuse in a national survey of adult men and women: Prevalence, characteristic and risk factors. *Child Abuse & Neglect, 14*, 198-28.

Finkelhor, D. (1979). *Sexually Victimized Children.* New York: Free Press.

Friedrich, W., Urquiza, A., & Bielke, R. (1986). Behavior problems in sexually abused young children. *Journal of Pediatric Psychology, 11*(1), 47-57.

Gill, M., & Tutty, L. (1997). Sexual identity issues for male survivors of childhood sexual abuse: A qualitative study. *Journal of Child Sexual Abuse, 6*(3), 31-47.

Goldman, J., & Padayachi, U. (2000). Some methodological problems in estimating incidence and prevalence in child sexual abuse research. *Journal of Sex Research, 37* (4), 305-315.

Goodman, L. Dutton, M., Harris, M. (1995). Episodically homeless women with serious mental illness: Prevalence of physical and sexual assault. *American Journal of Orthopsychiatry, 65*, (4), 468-478.

Haugaard, J.J. (2000). The challenge of defining child sexual abuse. *American Psychologist, 55*(9), 1036-1039.

Kiser, L., Heston, J., Millsap, P., Pruitt, D. (1991). Physical and sexual abuse in childhood: Relationship with posttraumatic stress disorder. *Journal of the American Academy of Child Adolescent Psychiatry*, *30*(5), 776-783.

Leitenberg, H., Greenwald, E. & Cado, S. (1992). A retrospective study of long-term methods of having been sexually abused during childhood. *Child Abuse & Neglect*,*16*, 339-407.

Perrott, K., Morris, E., Martin, J., & Romans, S. (1998). Cognitive coping skills of women sexually abused in childhood: A qualitative study. *Child Abuse & Neglect*, *22*(11), 1135-1149.

Peters, D., & Range, L. (1995). Childhood sexual abuse and current suicidality in college women and men. *Child Abuse & Neglect*,*12*, 201-208.

Rind, B., Tromovitch, P., & Bauserman, R. (1998). A meta-analytic examination of assumed properties of child sexual abuse using college samples. *Psychological Bulletin*, *124* (1), 22-53.

Rimza, M. E., Berg, R. & Locke, C. (1988). Sexual abuse: somatic and emotional reactions. *Child Abuse & Neglect*, *12*, 201-208.

Russell, D. E. H. (1983). The incidence and prevalence of intrafamilial and extrafamilial sexual abuse of female children. *Child Abuse & Neglect*, *7*, 133-146.

Russell, D. E. H. (1984). *Sexual exploitation: Rape, child sexual abuse and workplace harassment*. Beverly Hills, CA: Sage.

Saywitz, K. J., Mannarino, A. P., Berliner, L., & Cohen, J. A. (2000). Treatment for sexually abused children and adolescents. *American Psychologist*,*55* (9), 1040-1049.

Schaaf, K., & McCanne, T. (1998). Relationship of childhood sexual, physical and combined sexual and physical abuse to adult victimization and posttraumatic stress disorder. *Child Abuse & Neglect*, *22*(11), 1119-1133.

Steiger, H. & Stanko, M. (1990). Sexual trauma among eating-disordered, psychiatric and normal female groups. Comparison of prevalence and defense styles. *Journal of Interpersonal Violence*, *591*, 74-86.

Wells, R., McCann, J., Adams J., Voris, J., & Ensign, J. (1995). Emotional, behavioral and physical symptoms reported by parents of sexually abused, nonabused and allegedly abused prepubescent females. *Child Abuse & Neglect*, *19*(2), 155-163.

Wheeler, B. & Walton, E. (1987). Personality disturbances of adult incest victims. *The Journal of Contemporary Social Work*, December, 597-602.

White, S., Halpin, B., Strom, G. & Cental, G. (1988). Behavioral comparisons of young sexually abused, neglected and nonreferred children. *Clinical Child Psychology*, *17*(1), 563-71.

Wolfe, V., Gentile, C., & Wolfe, D. (1989). The impact of sexual abuse on children: a PTSD formulation. *Behavior Therapy*, *20*, 215-228.

Wozencraft, T., Wagner, W. & Pellegrin, A. (1991). Depression and suicidal ideation in sexually abused children. *Child Abuse & Neglect*, *15*, 505-511.

Blood Types and Athletic Performance

Gordon W. Russell
University of Lethbridge, Canada

Masao Ohmura
Nihon University, Tokyo

The participants were Japanese professional baseball players (N = 734) whose playing positions and team performance records were analyzed with respect to blood types. Chi-square analyses yielded non-significant relationships on playing positions and team performance. Players who since 1950 led their league in a number of season performance categories, e.g., batting, home runs, runs batted in, pitching, earned runs average, and most valuable player, were compared to those on the 1996 rosters. They did not differ from their 1996 counterparts. The distribution of blood groups of winning managers/coaches did not differ from those leading less successful teams.

The impression conveyed by Western news reports is that Japanese popular culture is awash with beliefs that blood types are predictive of a variety of human behaviors. For example, during World War II, army and navy battle groups were said to have been formed on the basis of blood types (Sullivan, 1996). Blood types also figure prominently in affairs of the heart whereby prospective mates are sought on the basis of blood types (Treen & Hoshiai, 1985). Additionally, some members of the business community subscribe to a belief that blood type is related to managerial success (Coscarelli, Stepp & Lyerla, 1989). For example, job advertisements have been known to specify that blood Types A and O are being sought, Type B's need not apply! (Treen & Hoshiai, 1985). Survey data indicates that these beliefs are widespread, more strongly held by females and related to beliefs in paranormal phenomena, e.g., astrology, palmistry, fortune telling (Nishizawa, 2000).

The literature is dominated by the works of Nomi and his son (e.g., Nomi, 1985). By one account they have written 25 books on the topic (Sullivan, 1996), by another 35 books (Treen & Hoshiai, 1985) and sold over six million copies. However, their conclusions of a relationship between blood types and various behaviors have been challenged. For example, a series of investigations by Ohmura (1994) has failed to find evidence supporting a relationship (also see, e.g., Coscarelli et al., 1989; Murata, 1994; Thompson, 1936). Nevertheless, other studies have found some support for a relationship between blood types and personality (e.g., Cattell, Boutourline Young & Hundleby, 1964; Eysenck, 1982;

Author info: Correspondence should be sent to: Prof. Gordon Russell,
1 McKenzie Crescent, Sidney, BC Canada, V8L 5Y7, e-mail: russell@uleth.ca.

North American Journal of Psychology, 2002, Vol. 4, No. 1, 161-164.

Furukawa, 1930; Jogawar, 1983) as well as season of birth (Gupta, 1992) and suicidal behaviors (Zonda, Csiszer & Lester, 1999).

Such empirical evidence as we have suggests that Type O coaches and managers would predictably be found at the helm of more successful teams. That is, analyses of questionnaires administered to 20,000 Japanese prompted Suzuki to characterize Type O's as tending "... to take leadership roles because of their decision-making capability, strong will, and realistic outlook" (cited in Coscarelli et al., 1989). In summarizing Nomi's writings, Sullivan describes Type O's as "... powerful leaders, goal-oriented, enthusiastic, optimistic and good at business" (Sullivan, 1996, p. 5). The present study then examines the distribution of blood groups between the successful and the less-than-successful baseball coaches/managers.

Performance is a variable that is central to the interests of sport investigators (see e.g., Russell, 2001). In this regard, we tested the supposition that blood types are associated with playing positions on the diamond. Also, we tested predictions that the performance of teams and that of historical leaders in a number of major performance categories is related to blood groups.

METHOD

Participants

The participants (N = 734) were professional baseball players in Japan's six-team, Central and Pacific Leagues for whom blood types were available. Non-native Japanese (approximately 6% of all team rosters) were excluded from the analyses.

Procedure

Individual and team performance measures for the 1996 season were taken from the *Japan pro baseball handbook & media guide* (Graczyk, 1997). Playing positions and measures of managerial success were also obtained from the handbook. The individual performance records and blood types of season leaders in both leagues covered the period from 1950 to 1996 and included the leaders in batting, home runs, runs batted in (rbi), pitching, earned runs [era] and the most valuable player [mvp]. Winners in multiple years were counted only once.

RESULTS AND DISCUSSION

The distribution of the four blood types within the sample of ballplayers was as follows: A = 37.74%, O = 32.70%, B = 20.71%, and AB = 8.85%. As presented in Table 1, the distribution of players' blood types across playing positions, i.e., pitchers, catchers, infielders, outfielders, in each league and overall did not differ from chance expectations. Thus, no one type has favored ballplayers at particular positions as they develop their skills and rise through the ranks of organized baseball. Neither would it seem that coaches have been actively recruiting for positions on the basis of blood types.

TABLE 1 Blood Types A, B, O, & AB in Relation to Player Position and Baseball Performance

	N	Chi-Square	*df*	*p*
Playing Position				
Central league	368	3.88	9	.92
Pacific league	366	6.33	9	.71
Combined	734	7.70	9	.57
Performance				
Managers/coaches (winning vs. losing)	107	1.26	3	.74
Teams (winning vs. losing)	245	3.02	3	.39
Leaders (1950 - 1996): batting	774	3.62	3	.31
home runs	765	1.86	3	.60
rbi	764	3.86	3	.28
pitching	375	3.58	3	.31
era	766	2.34	3	.50
mvp	782	4.18	3	.24

The prediction that effective leadership is in some way related to blood Type O (cf. Suzuki, cited in Corscarelli et al., 1989) failed for lack of support. The coaching staffs and managers who guided their teams to either the top or bottom halves of their respective leagues did not differ in regard to the distribution of blood types.

The two teams in both leagues with the best win/loss records for 1996 were compared to those finishing in the last two positions of the leagues. The rosters of teams with winning seasons did not differ in the distribution of the four blood groups from teams that had losing seasons.

The blood groups of season leaders in six performance categories over a span of nearly half a century were compared to those players on the team rosters for 1996. Note, however, that the comparisons would ideally be made with same-year cohorts. In each category, blood type was unrelated to performance excellence. In the case of pitching prowess, the historical leaders were compared to the 1996 crop of pitchers. Again, there was a non-significant relationship. Winning teams and outstanding athletes have blood types distributed not unlike those who are less successful. Athletic success appears unrelated to blood groups. Individual performance criteria in baseball are seemingly better predicted by other factors, e.g., physical attributes, age (Russell, 2001). Our results speak clearly to a number of plausible notions regarding blood groups and performance in baseball. Notwithstanding the popularity of blood type beliefs in Japanese culture, support for a series of predictions tested at the professional level of the sport was not forthcoming.

REFERENCES

Cattell, R. B., Boutourline Young, H., & Hundleby, J. B. (1964). Blood groups and personality traits. *American Journal of Human Genetics, 16,* 397-402.

Coscarelli, W. C., Stepp, S. L., & Lyerla, R. (1989). Relationship of blood type with decision-making style and personality type. *Mankind Quarterly, 29,* 307-327.

Eysenck, H. J. (1982). The biological basis of cross-cultural differences in personality: Blood group antigens. *Psychological Reports, 51,* 531-540.

Furukawa, T. (1930). A study of temperament and blood groups. *The Journal of Social Psychology, 1,* 494-509.

Graczyk, W. S. (1997). *Japan pro baseball fan handbook & media guide.* Tokyo: Yohan.

Gupta, S. (1992). Season of birth in relation to personality and blood groups. *Personality and Individual Differences, 13,* 631-633.

Jogawar, V. V. (1983). Personality correlates of human blood groups. *Personality and Individual Differences, 4,* 215-216.

Murata, K. (1994). Why do they believe blood groups correlate with personality differences?: An introduction to experimental research on social cognition. *Hitotsubashi ronso, 111,* 119-135 (Japanese text).

Nishizawa, S. (2000). Psychic well-being, religion, logos intuition and rational emotive therapy. *Proceedings of the 58th Annual Convention of the International Council of Psychologists,* Padu, Italy.

Nomi, M. (1985). *Ketsuekigata to seikaku hando bookku* [Blood type and personality handbook]. Tokyo: Sei Shun Shuppan Co.

Ohmura, M. (1994). *Ketsuekigata to seikaku* [Blood types and personality]. Tokyo: Fukumura Publisher.

Russell, G. W. (2001). *Sport science secrets: From myth to facts.* Victoria, Canada: Trafford.

Sullivan, K. (1996, January 1). A or O? B or AB? In Japan, blood type is the be all and end all. *International Herald Tribune,* pp. 1, 5.

Thompson, G. N. Jr. (1936). Blood type as related to intelligence, emotions, and personality. *Journal of Applied Psychology, 20,* 785-789.

Treen, J., & Hoshiai, Y. (1985, April 1). 'Typecasting' - by blood. *Newsweek.*

Zonda, T., Csiszer, N., & Lester, D. (1999). Blood groups among attempted and completed suicides. *European Journal of Psychiatry, 13,* 58-60.

Authors' Note: The writers wish to thank Dr. Mark L. Sandilands, Prof. Satoru Nishizawa, Dr. Hal Weaver and Dan Latham for their assistance.

Predicting Magnitude Estimates of Drink Strength

Sharolyn R. Higgs, Stuart J. McKelvie and Lionel G. Standing
Bishop's University

Thirty undergraduates gave direct magnitude estimates for the strength of vodka/water and of grape juice/water mixtures, using concentrations from 10% to 100%. Correlations between actual and judged strength were very high (.990 for vodka and .997 for grape juice), but corresponding exponents for Stevens' power law were below unity (.73, .83), indicating underestimation at the highest concentrations. Vodka and grape juice accuracy judgments were separately regressed on a number of subject variables (including alcohol consumption, smoking, and exercise), which did not predict them significantly. However, on the basis of *differences* between grape juice and vodka accuracy, the magnitude estimates of vodka concentration were relatively better as the number of years of drinking experience increased. Suggestions are made for future research.

Ninety percent of university students drink or have tried alcohol (Engs & Hanson, 1989; Svenson, Jarvis, & Campbell, 1994), and 30% to 50% engage in binge drinking (Borsari & Carey, 1999). Therefore, an important question is: How accurately can they judge drink strength?

According to Stevens (1975), the subjective magnitude of a stimulus follows the power law $\varphi = k\ \Phi$, where φ is subjective magnitude, Φ is physical magnitude, n is the exponent, or the slope of the log-log plot relating ϕ to Φ, and k is a scale constant. For taste, Stevens (1975, p. 15) reports exponents (n) of 1.3 for sucrose and 1.4 for salt, indicating progressive overestimation at higher magnitudes, and 0.8 for saccharine, indicating underestimation at the top.

In the only apparent study of alcohol, Standing and Blackburn (1995) found that the correlation between actual and judged strength of white rum was .99, over ten rum/water concentrations from 10% to 100%. The slope of the regression line for actual and judged concentrations was 1.03, showing "neither overestimation or under- estimation." (p. 365). However, subsequent calculation from log-log coordinates showed that n was 1.25, indicating overestimation at higher strengths. The purpose of the present experiment was to replicate these results with vodka/ water mixtures, and to follow up Standing and Blackburn's suggestion that drinking experience may predict accuracy. Participants were also run in a non-alcoholic (grape juice) control condition. It was predicted

---

Author info: Correspondence should be sent to: Dr. Stuart McKelvie, Psych. Dept., Bishop's University, Lennoxville, Quebec, Canada. J1M 1Z7.
North American Journal of Psychology, 2002, Vol. 4, No. 1, 165-170.
© NAJP

that only vodka accuracy would increase with more drinking experience (operationalized as number of years drinking alcohol and amount consumed per week).

Some subject factors influence gustatatory *sensitivity*. For example, detection thresholds are higher with age, male gender, smoking, and hypoglycemia (Cowart, 1981; Pasquali, 1997). However, subject variable predictors of *magnitude estimation* accuracy have not been investigated. In addition to drinking experience, the present study examined other factors, particularly smoking and exercise. From the detection research, it was expected that both vodka and grape juice accuracy would decline for participants who smoked more. A novel feature of this study is the inclusion of exercise, whose effect on gustatory sensitivity has not been investigated. We predicted that accuracy would decline for people who exercised less, based on the premise that less active people may have less efficient blood flow, thereby reducing sensory functioning.

METHOD

Participants

Fifty-four undergraduates from a variety of disciplines completed a survey of drinking[1], smoking, exercise and other variables. To avoid restriction of range problems, the 30 selected experimental participants (18 women, 12 men; mean age 21.6 yr.) included people who reported high, medium and low levels of drinking, smoking, and exercise.

Materials and Procedure

The survey covered years of drinking experience, alcohol consumed per week[2] (sum of beer, wine and hard liquor units), cigarettes smoked per week, exercise hours per week, and demographic variables (age, gender, home province, rural/urban background, reported mean grade, year of study, on/off campus living). Students were recruited in a variety of campus locations (e.g., sports centre, pub) to ensure that the sample would reflect different levels of drinking, smoking and exercise.

About three weeks after completing the survey, the 30 test students participated in a direct magnitude estimation task in which they gave 20 absolute percentage judgments of drink strength. The drinks used were *Prince Igor's Russian Vodka* (40% by volume) and *Minute Maid's Frozen Purple Grape Juice Concentrate*. Vodka/water mixtures (5ml) were given on 10 trials and grape juice/water mixtures (5ml) were given on 10 trials, with half of the participants receiving vodka first and half receiving juice first. Each mixture was presented in separate 30ml plastic shot glasses. Concentrations ranged from 10% to 100% in steps of 10%. Test order was governed by a 10 x 10 Latin square design. The 50% modulus was presented before each trial. Following each magnitude estimation, participants expectorated, then rinsed their mouths with demineralized water. Participants initially completed an informed consent form and were debriefed after making their judgments.

RESULTS AND DISCUSSION

Replicating Standing and Blackburn (1995)'s correlation of .99 for rum, the correlation for vodka between the 10 physical magnitudes and the 10 mean subjective magnitudes averaged over the 30 participants was .990. The correlation for grape juice was .997. As concentration increased, subjective magnitude increased (see Figure 1). The slopes for the least-squares regression lines for the raw data were 0.790 (vodka) and 0.903 (juice), and the exponents (n) for Stevens' power law, calculated as the slope of the corresponding log-log functions, were 0.727 and 0.826. These values are lower than Standing and Blackburn's estimates of 1.03 (raw data slope) and 1.25 (log-log slope) for rum. Thus, whereas the strength of rum was slightly *over*estimated at higher concentrations (a power function that is slightly concave upwards), the strength of vodka (and, to a lesser extent, grape juice) was slightly *under*estimated (power functions that are slightly concave downwards). Indeed, single-sample t-tests for each mean subjective magnitude compared to the correct physical magnitudes showed that vodka was overestimated with 10% concentration [M = 18.67, SD = 14.08, $t_{(29)} = 3.37$, $p < .01$], but underestimated at both 90% [M = 78.33, SD = 14.16, $t_{(29)} = 4.51$, $p < .01$] and 100% [M = 87.33, SD = 16.60, $t_{(29)} = 4.18$, $p < .01$]. Also, juice was underestimated at 100% [M = 93.67, SD = 10.33, $t_{(29)} = 3.36$, $p < .01$].

The exponent for grape juice (.826) is similar to Stevens' estimate of .80 for saccharine. The difference between the exponents for rum and for vodka may be due the fact that rum has a more distinctive flavor than vodka. If this is the case, other spirits with a distinctive flavor (e.g. scotch) would be overestimated at higher strengths and other spirits with a less distinctive flavor (e.g. gin) would be underestimated. In addition, although we followed Standing and Blackburn's procedure, Stevens (1975, p. 15) notes that exponents for the same dimension (e.g., brightness) may vary with small differences in experimental setting.

Participant accuracy was estimated in three ways: the *correlation* between physical and subjective magnitude across the 10 concentrations, the *number of correct judgments* (maximum = 10), and the *mean absolute difference* between physical and subjective magnitudes over the 10 concentrations. The mean correlations averaged over the 30 participants were fairly high (M = .783, SD = .178 for vodka; M = .883, SD = .154 for juice), but the mean number correct was low (M = 2.62, SD = 1.45 for vodka; M = 4.73, SD = 2.66 for juice), and the mean difference scores were 14.17% (SD = 6.18) for vodka and 8.63% (SD = 6.46) for juice. In each case, performance was better for juice than for vodka, $t_{(29)} = 2.50, 3.91$, and 3.74, $ps < .05$. However, although participants tasted very small amounts (5 ml), and may not have looked at the liquids, this improvement may have been partly due to the color of the grape juice mixture, which was more saturated at higher concentrations. This confounding variable compromises the conclusion that taste accuracy was higher for juice than for vodka.

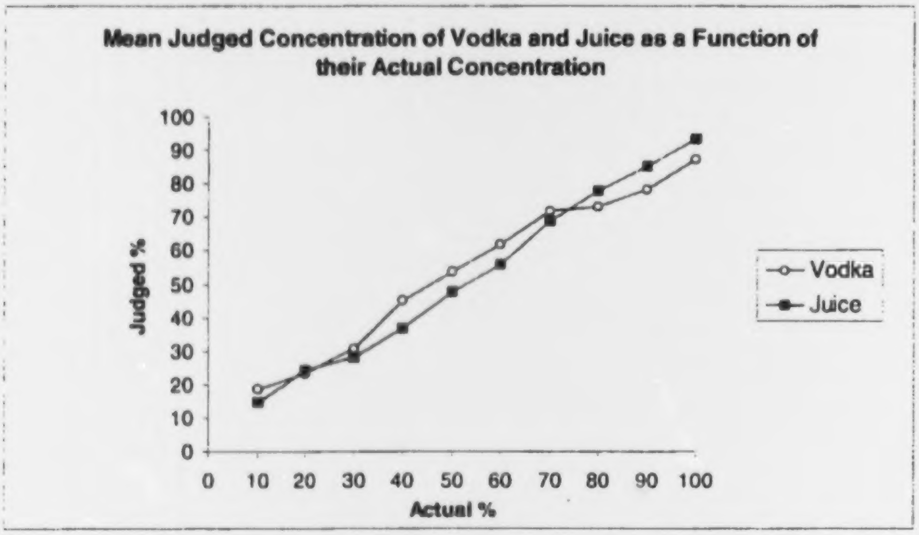

Figure 1. Mean Judged Concentration of Vodka and Juice as a Function of their Actual Concentration

Pearson product-moment correlation coefficients for the 30 participants ranged from .236 to .964 for vodka and from .251 to 1 for grape juice, ranges that are wider than Standing and Blackburn's (1995) range of .84 to .98 for rum. Absolute accuracy ranged from 0% to 6% for vodka and from 0% to 10% for juice, while mean difference scores ranged from 6% to 28% for vodka and from 0% to 31% for juice.

To predict these variations from participant to participant, each measure of magnitude estimation accuracy was considered separately as a dependent variable and regressed on the survey variables. None of the predictors was significantly correlated with any of the three measures. However, zero-order correlation coefficients might reflect restriction of range on the variables. Although absolute accuracy was very low and near floor, mean difference scores had a healthy range and correlation scores had a very wide range. Moreover, the number of cigarettes smoked per week ranged from 0 to 200. Therefore, the zero-order correlations between smoking and accuracy were assumed to be valid.

They contrast with the positive relationships observed between smoking and taste absolute thresholds (Cowart, 1981), but they support Stevens's (1961) contention that the steepness of the power function (the scaling problem) is independent of "resolving power" (the detection problem). Future research might include smoking as a predictor of both detection and of magnitude estimation accuracy for the same drinks.

Drinking experience was not related to juice or to vodka accuracy with separate analyses. However, because it had been hypothesized to predict vodka but not juice judgments, a fourth dependent variable was calculated: the *difference* ('juice–vodka') between vodka and juice accuracy for each person on each accuracy measure. A positive value indicates better juice accuracy for *correlations* and for *number correct*, whereas a negative value indicates better juice accuracy for *difference scores*. For correlations, $M = .100$, $SD = .219$, range = -.627 to .764), for number correct, $M = 2.10$, $SD = 2.94$, range = -3 to 9, and for difference scores, $M = -5.53$, $SD = 8.10$, range = -28 to 16.

For the new dependent variable, two correlations were significant: -.367, $p < .05$ for number of years of drinking experience and correlation score, and .376, $p < .05$ for years drinking experience and difference score. Marginally significant ($p < .10$) correlations were also found between number of units of alcohol drunk per week and both juice-vodka absolute accuracy (-.311) and difference scores (.323), and between number of cigarettes smoked and juice-vodka absolute accuracy scores (-.34). With smoking was as a binary variable (smoke, don't smoke), the correlation with juice-vodka absolute accuracy was -.36, $p < .05$.

In the multiple regression analyses, the two significant relationships between years of drinking and juice-vodka accuracy remained significant. For the correlation score, $R^2 = .135$, $b = -0.025$; for the difference score, $R^2 = .142$, $b = 0.938$. That is, as the number of years of drinking increased, the accuracy advantage of juice over vodka declined, indicating relatively better magnitude judgments of vodka concentration. In addition, for juice-vodka absolute accuracy, smoking as a binary variable was the only significant predictor, $R^2 = .132$, $b = -2.625$. Unfortunately, this intriguing finding that the ability to judge concentration of vodka relative to juice was better for smokers than for non-smokers only occurred with one of the three measures of accuracy.

A novel feature of this experiment was the inclusion of exercise as a predictor. Contrary to our intuition that less exercise would be related to less accuracy because of decreased blood flow, the correlations were very close to zero (-.07, .11, and -.02 for the correlation, absolute, and mean difference measures of accuracy respectively). It is unlikely that they reflect restriction of range, because the number of exercise hours per week ranged from 1 to 50. Perhaps exercise, like smoking, predicts detection but not magnitude estimation.

Although care was taken to ensure that the measured variables were not restricted in range, there were only 30 undergraduate participants. Because drinking is a popular pastime in this population, results with students are of interest, but the findings cannot be generalized beyond them. Replication with more men and women over a wider age range from the general population would be valuable.

CONCLUSION

Participants judged rather well how concentrated an alcoholic drink (vodka) was, although they tended to underestimate at the highest levels. This could have serious consequences for over-consumption if drinks are high strength. Judgments of grape juice concentration were more accurate than judgments of vodka, but the advantage was smaller for people who had more experience drinking alcohol. Future research should eliminate the confounding variable of drink color, examine further the three different measures of accuracy, and explore drinking experience, smoking, and exercise as predictors of both magnitude estimation and detection for various alcoholic beverages.

REFERENCES

Borsari, B. E., & Carey, K. B. (1999). Understanding fraternity drinking: Five recurring themes in the literature, 1980-1998. *The Journal of American College Health, 48*, 30-37.

Cowart, B. (1981). Development of taste perception in humans: Sensitivity and preference throughout the life span. *Psychological Bulletin, 90*, 43-73.

Engs, R., & Hanson, D. J. (1989). Reactance theory: A test with collegiate drinking. *Psychological Reports, 64*, 1083-1086.

Higgs, S.R., McKelvie, S. J., & Standing, L. G. (2001). Students' reports of athletic involvement as predictors of drinking: A pilot study. *Psychological Reports, 89*, 487-488.

Pasquali, B. (1997). Menstrual phase, history of smoking, and taste discrimination in young women. *Perceptual and Motor Skills, 84*, 1243-1246.

Standing, L., & Blackburn, A. (1995). Can we judge the strength of a drink? *Perceptual and Motor Skills, 81*, 365-366.

Stevens, S. S. (1975). *Psychophysics: Introduction to its perceptual, neural, and social aspects.* New York, Wiley.

Stevens, S. S. (1961). To honor Fechner and repeal his law. *Science, 133,* 80-86.

Svenson, L., Jarvis, G., & Campbell, R. (1994). Gender and age differences in the drinking behavior of university students. *Psychological Reports, 75,* 395-402.

Footnotes:

[1] There is a separate report of the drinking scores for all participants (Higgs, McKelvie, & Standing, 2001).

[2] We thank Ruth Engs and Brian Borsari for questionnaire advice.

How to Submit Papers to *NA*JP

The submission of good papers is very important to any psychology journal. We hope that you will support our views about what a psychology journal should be like by submitting papers now. We will try to mail a decision on your paper within 5 weeks of receipt by us. We will publish your paper in an average of 5 months of receipt of a final accepted version.

Submit 1 paper copy of each manuscript prepared according to APA guidelines. Also include a diskette with your manuscript or (preferably) e-mail the file to NAJP@mpinet.net. Include 2 large stamped envelopes so we can send your manuscript out for review. We encourage you to have your paper read by peers before submission.

If the final paper is accepted, please send a check or purchase order for the amount specified with your final draft (billed at $25 dollars per typeset page in units of two pages).

Submission of a manuscript to the Journal is a contract to publish with the Journal if the paper is accepted. Because of our rapid editorial feedback, simultaneous submission to other journals is not allowed. We will offer brief comments on the suitability of a topic for publication in *NA*JP before a submission.

Send all correspondence to: *North American* Journal of Psychology, 240 Harbor Dr., Winter Garden, FL 34787, phone 407-877-8364.

SUBSCRIPTIONS

All subscriptions are on a calendar year basis only. If you order late in the year, the earlier issues will be sent to you. Institutional rate, $65. Individual rates, paid by check, $30. Add $7 for Canadian subscriptions. Add $8 for subscriptions outside North America. Phone orders accepted with Visa, MasterCard or American Express. Call 407-877-8364. The *North American* Journal of Psychology is now published semi-annually in June & December.

Visa ☐ AMEX ☐ MC ☐

Card # Exp. Date

Name

Address

City State Zip

A Summary of Our Philosophy

The ***North American*** **Journal of Psychology** publishes scientific papers that are of general interest to psychologists and other social scientists. Papers may encompass a wide variety of topics, methods, and strategies. It is our belief that many good scientific papers are rejected for invalid reasons. Some journals reject up to 90% of their submissions because they don't have space. Furthermore, studies of the review process strongly suggest that not all of the lucky 10% are better papers than those that get rejected. In our view many of those rejected papers are meritorious with need for only minor revisions. Perhaps our philosophy can be summed up this way: we are looking for ways to strengthen good papers rather than searching for weak excuses to reject them.

North American
Journal of Psychology

Vol. 4, No.2 **2002**

Advisory Editorial Board

ADRIAN FURNHAM
Psychology
University College London

DOUGLAS JACKSON
Psychology
University of Western Ontario

JAMES M. JONES
Psychology
University of Delaware

HEIDI KELLER
Psychology
University of Osnabrueck

MANUEL LONDON
Psychology
SUNY, Stony Brook

KAREN LYNESS
Management
Baruch College, CUNY

ROBERT PASNAK
Psychology
George Mason University

RICHARD RYAN
Psychology
University of Rochester

PHILIP ZIMBARDO
Psychology
Stanford University

Editor
LYNN E. McCUTCHEON

How to Subscribe to The *North American* Journal of Psychology

All subscriptions are on a calendar year basis only. If you order late in the year, the earlier issue will be sent to you. Institutional rate: $65. Individual rate: $30. Add $7 for Canadian subscriptions. Add $8 for subscriptions outside North America. *North American* Journal of Psychology is published in March, June and December.

Now visit our website at http://najp.8m.com

North American Journal of Psychology

July, 2002 ISSN 1527-7143 Vol. 4, No.2

An Interview With Arnold A. Lazarus 171
Arnold A. Lazarus & Michael F. Shaughnessy

Depression, Drug Use, and Gender Differences Among Students at a Religious University 183
Herbert W. Helm, Jr., Mark D. Boward, Duane C. McBride, & Rachael I. Del Rio

Participation Motives in Predicting Sport Dropout Among Australian Youth Gymnasts 199
Todd A. Ryska, David Hohensee, Dean Cooley, & Christine Jones

The Internal Consistency, Reliability & Construct Validity of the German Translation of the Oxford Happiness Inventory 211
Christopher Alan Lewis, Leslie J. Francis, & Hans-Georg Ziebertz

Body-image Assessment: A Review and Evaluation of a New Computer-aided Measurement Technique 221
Salvatore Cullari, Michelle Vosburgh, Amber Shotwell, Julien Inzodda, & Wendi Davenport

Personality and the Aesthetics of Composition: A Study of Mondrian and Hirst 233
Adrian Furnham & Sumedh Rao

A Short Scale of Family Atmosphere (SOFA): Development and Psychometric Evaluation 243
Geoffrey N. Molloy & Julie F. Pallant

The Sport Widow Hypothesis: A Research Note 249
Gordon W. Russell & Robert L. Arms

Effect of Different Normal/bizarre Image Combinations on Pair-cued Associated Word Recall 253
Alfredo Campos, Maria Jose Perez-Fabello, & Maria Calado

Explaining Punitiveness: Right-wing Authoritarianism and Social Dominance 263
Jason S. Capps

Status, Apparel and Touch: Their Joint Effects on Compliance to a Request 279
Nicolas Gueguen

Do His Family and Friends Like Me? Predictors of Infidelity in Intimate Relationships 287
Ann Zak, Colleen Coulter, Sabrina Giglio, Jennifer Hall, Stephanie Sanford, & Nancy Pellowski

An Interview with Mark Ylvisaker about Students with Traumatic Brain Injury 291
Mark Ylvisaker, Michael F. Shaughnessy, & Dan Greathouse

Delivering Human Services to Native Americans with Disabilities: Cultural Variables & Service Recommendations 309
Deborah Rowley & Ruth Anne Rehfeldt

On Differentiating Major Depression from Chronic sadness: A Commentary 317
Salvatore Cullari

Editor's Comments

Thanks again to all the reviewers, all of whom did a fine job of critiquing papers promptly and with great insight. Without our reviewers, this journal would simply not have the quality it has today. Beginning with this issue, I want to recognize three reviewers by name who have done a fine job of reviewing two or more papers for NAJP. For this issue a tip of the editor's cap goes to Wayne Wilson, David L. Johnson, and Tammy Chapman.

I am heartened by the appearance of an intriguing article by Yale's Robert Sternberg (Dr. Sternberg was interviewed for NAJP last year by Mike Shaughnessy. Sorry - couldn't resist the opportunity to "plug" NAJP). The point of Sternberg's comments was that some reviewers go out of their way to make nasty, sarcastic, or otherwise destructive comments about papers. Sternberg feels that such comments are inappropriate and should have no place in the review process. I agree, and I want to add that I am proud of our NAJP reviewers for their professionalism in this regard. *Constructive* criticism is essential to the quality of a scientific journal. Our reviewers have raised good questions about methodology, writing style, and issues relating to clarity *without* getting nasty.

In the last issue I applauded the willingness of certain APA editors to make their journals more "friendly." I want to continue to encourage other editors to do the same.

Lynn E. McCutcheon, editor

An Interview With Arnold A. Lazarus

Arnold A. Lazarus
Rutgers University

(interviewed on behalf of NAJP by)
Michael F. Shaughnessy
Eastern New Mexico University

NAJP: The last twenty-five years have seen a proliferation of computerized assessment devices- the M.M.P.I., the 16 P.F. and various other instruments that are now computer scored and analyzed. How has this affected the way you practice and how has it affected psychology in general?

AL: At the beginning of my career, I was heavily involved in testing. Nearly all my clients filled out various questionnaires (including the MMPI and a range of personality, neuroticism, and adjustment inventories). Frankly, it did not take me very long to realize that I did not find these psychometric instruments helpful or useful in treatment planning. In graduate school, I was required to study projective tests -- primarily the Rorschach, Thematic Apperception Test and various sentence completion scales. Here again, in my clinical practice, I found these devices marginally helpful. Instead, I concluded that careful history taking, meticulous observation, and corroboration or refutation from the client's significant others paved the way to discovering information that could be put to good use in therapy. I do not think that measuring instruments point the way to appropriate therapeutic methods and strategies. Frankly, I question the true validity of most of these scales. I have tried out several computerized instruments, but it is my contention that BASIC I.D. and Second-Order BASIC I.D. assessments will shed more light on the selection of effective treatment protocols than the aforementioned tests.

NAJP: Over the last twenty years H.M.O.s have literally terrorized the field of medicine and psychology. How can therapists and psychologists cope with this insidious " big brother "? How will this affect the way we do business in the future?

Author info: Correspondence should be sent to: Dr. Michael Shaughnessy, Psychology Dept., Eastern New Mexico U., Portales, NM 88130

North American Journal of Psychology, 2002, Vol. 4, No. 2, 171-182.
© NAJP

AL: Many of the therapists I know have nothing whatsoever to do with insurance companies or managed care. Instead of all the paperwork and haggling involved, they prefer to charge less, and they provide their clients with a receipt and diagnosis after each session, which the client can turn into his or her insurance company and usually, receive some reimbursement.

NAJP: What is your assessment of the current state of the art of medication as an adjunct to psychotherapy?

AL: As we all know, there is a raging controversy as to the value and effectiveness of medication for many conditions. While I think it is true that the drug companies have oversold the putative benefits of their wares, there are undoubtedly certain clients who require appropriate medication. Perhaps one of the most obvious instances is that of bipolar disorders, in which the use of medication such as lithium is essential. Recently, I was treating a woman who suffered from depression for many years but who refused to take antidepressants. She received psychotherapy from some skillful professionals, all to no avail. I too, was unsuccessful until I managed to persuade her to see a psychiatrist who prescribed a SSRI. Within a few weeks she was remarkably less depressed. Nevertheless, the “total cure” did not come out of a bottle because when her depression abated, we were able to tackle other psychosocial issues that could not be addressed when she was in the throes of extreme depression. There have been many cases in which medication as an adjunct to psychotherapy has proved invaluable.

NAJP: Your ideas of "bridging and tracking" have been well received. What other new concepts and constructs have you developed recently?

AL: My approach to therapy strives to remain current with cutting edge findings across disciplines. Thus, in keeping with the need to find “treatments of choice,” I have endeavored to remain *au courant* with the research on *empirically supported methods*, and to apply them whenever feasible. At this juncture, new concepts and constructs seem less important than consolidation and verification with a view to determining which methods to use with whom, and for which particular conditions.

NAJP: The last twenty years have seen a growth in relationship and marital counseling. Why do men and women have such a difficult time understanding and relating to one another, and how can counselors help?

AL: As I had written in my book *Marital Myths* and in my updated *Marital Myths Revisited*, "most people do not know how to be married." There are no mandatory courses on marriage. It would be helpful if people who end up with a marriage license have at least a modicum of expertise in marriage. Instead, many people marry for the wrong reasons, enter their conjugal relationships with ridiculous expectations, and they often treat one another abominably in the name of "spontaneity," "being oneself" and "letting one's hair down." Surely everybody knows that it is a very bad idea to criticize one's spouse, to dredge up past resentments, to mind read, to interrogate their partner, and to become a control freak. So why do they do these things? Sometimes it is because they are sick people. Often it is because they are truly incompatible. Some individuals do not know how to express appreciation and caring. And, by the way, this has little to do with gender. I have treated many gay and lesbian couples who make the same mistakes.

NAJP: Who has most influenced you and your work, and why?

AL: This is a difficult question for me to answer because I have learned so much from so many different people, too numerous to mention – from their books and other writings, workshops, seminars, and informal discussions. Yet some of the most powerful influences have been negative. I suffered at the hands of many inept professors and supervisors during my career and the influence they had on me was on what **not** to do and how **not** to be. Reverse role modeling you might say. On the positive side, a 12th grade high school teacher inspired me to want to learn and he instilled in me the desire to obtain an education. Guest lectures by the late Joseph Wolpe, M.D., initially stimulated me to be wary of the psychoanalytic concepts that dominated my undergraduate and graduate experiences, although he and I later parted company on theoretical grounds (I came to regard him as too narrow and rigid). Starting in the 1960s I had long conversations with Aaron Beck and Albert Ellis, and I also sat in on their classes and participated with them on several panels. I have been influenced by many friends, colleagues, students, post-doctoral trainees, and, of course, numerous clients.

NAJP: As we enter the new millennium where are the fields of counseling and psychotherapy going and why?

AL: We seem to have two major trends emerging simultaneously. The one I like better is the emphasis on *treatments of choice* for specific conditions, tied to *empirically supported methods*. In concert with the foregoing, we have the development of *manualized treatment protocols*. When researchers spell out step-by-step precisely what they do when administering specific treatments, it becomes possible to test and replicate their results. Treatment fidelity has become most important. In other words, for anyone to claim that a method was ineffective, the first question to ask is whether it was properly carried out and correctly applied. Without the foregoing, we remain in the Dark Age of subjective beliefs about treatment efficacy, in which anecdotes hold sway over evidence.

In contrast to the foregoing, we also have a wave of New Age Therapies that are not properly tested but that nevertheless have wide appeal. These are tied into what has been called the "Power Therapies" in which ridiculous and extravagant claims beguile many counselors into spending thousands of dollars for training, whereupon they (naturally) are willing to defend their "new methods" almost to the death. When I have given talks and have mentioned, for instance, that the purveyors of EMDR (Eye Movement Desensitization and Reprocessing) and TFT (Thought Field Therapy) have over stated the putative benefits of these methods, their followers have almost wanted to come to blows with me.

It will be too bad for our profession and for the millions of people who receive counseling and therapy if the scientific and data driven approaches are eclipsed by the subjective and anecdotal methods alluded to above.

NAJP: What do you consider to be your most important books and why? What do you see as your greatest contribution to psychology and psychotherapy?

AL: I see my book *Behavior Therapy and Beyond* as one of my most important contributions. This book became a citation classic. It is arguably one of the first books on what has come to be known as *cognitive-behavior therapy,* and it humanized the field of behavior therapy. Three other books, *Multimodal Behavior Therapy, The Practice of Multimodal Therapy*, and *Brief But Comprehensive Psychotherapy: The Multimodal Way* encouraged clinicians to practice in a humane yet systematic and broad-spectrum manner. I see some of my main contributions as (1) Showing that breadth is usually more important than depth for effective psychotherapy. (2) Emphasizing the value of technical eclecticism, and stressing the dangers of theoretical integration. (3) Developing the "multimodal" BASIC I.D. model that spawned several unique assessment strategies (especially those we call bridging, tracking, and Second Order Assessments). A book I am presently working on (which should be available in August 2002) co-edited with Ofer Zur, Ph.D., *Dual Relationships and Psychotherapy*, might create quite a stir. Its 31 chapters argue against the sterile, distant, boundary-hugging trajectory that so many therapists adhere to.

NAJP: What do you see as the **current** critical issues in psychotherapy?

AL: I think the proliferation of therapists is singularly unfortunate and only spawns confusion. When I became interested in psychotherapy in the 1950s, the dominant orientation was psychoanalytic, and many felt that only physicians (especially psychiatrists) should perform therapy. Psychologists were seen as trained to do testing. Social workers administered no formal therapy but worked in agencies that attended to housing, adoption, and other social issues. Today, it seems that most psychiatrists deal solely with biological concerns, psychologists have sought to survive by going into research, schools, industry, or attaching themselves to primary care physicians, and even to dental offices. Indeed, I read one survey that suggested that the bulk of psychotherapy is conducted by social workers, and a variety of different counselors. The level and type of training is uneven and unequal. I know a "dance therapist" and an "art therapist" who both have lucrative practices – neither one of whom has proceeded beyond an undergraduate degree (not even in psychology). Unless a highly informed panel is able to develop a basic curriculum that *all* psychotherapists would be required to follow (as is true in the field of medicine) I fear that our discipline will be seen as a roost for charlatans.

NAJP: When you are successful with clients, what factors do you see as most operative? When you do not do as well, to what do you attribute the difficulty?

AL: Several studies have indicated that one of the most important attributes of a successful therapist is his or her ability to be *persuasive.* From my perspective, this implies that therapists succeed whenever they are able to influence a client to think and act differently. The factors that are inclined to enhance one's persuasive powers include rapport, mutual trust and respect, and the client's confidence that the counselor or therapist is experienced and knowledgeable. One of the major skills a therapist requires is to gain the client's willingness to try out new behaviors, to take various emotional risks, and to re-evaluate faulty cognitions that are brought to his or her attention. When I fail, it is primarily because I may have missed the important facets that called for change, or I may not have managed to convince the client that adherence to my recommendations will indeed prove beneficial. As we all know, some clients are just not ready to make certain changes and they require a long runway (extended rapport and trust building, plus repeated reassurance and illumination) before they become airborne (i.e. ready to take action).

NAJP: What are your current personal and professional activities?

AL: To answer this question I would like first to outline, very briefly, the main course of my professional journey, and then I will mention some changes that have accrued in recent years. In South Africa I had been in full-time private practice for some 5 years, but in 1963 I was a visiting assistant professor at Stanford University and soon became a full professor at Temple University Medical School. I then went to Yale University as a visiting professor and Director of Clinical Training, and in 1972 accepted an offer to join the Faculty of Arts and Sciences at Rutgers University as a distinguished professor of psychology. I always maintained a busy part-time private practice. On January 1, 1998, I retired from Rutgers and left Academia. This has given me the time to pursue other interests (such as literature, American history, and cosmology), to participate in legal cases as an expert witness, and to assail what I regard as needlessly restrictive boundaries that ethics boards and licensing committees impose on psychologists.

I meet with a hand-picked group of close colleagues in New York City once a week. We keep each other abreast of developments in the field and debate a host of stimulating issues, theories, and occurrences. I also meet with a group of post-doctoral fellows, most of whom are my former students. The foregoing activities keep me occupied, and I also see clients, sometimes in concert with my son, Clifford, a Ph.D. licensed psychologist, and my daughter-in-law Donna, a Licensed Clinical Social Worker.

NAJP: Looking back, of what accomplishments and awards are you most proud?

AL: I have always felt particular satisfaction and joy when I have managed to render significant help to clients who had failed to respond to the ministrations of other therapists – some of whom are regarded as luminaries in the field. I take pride in having been of value to people whose IQs are far higher than mine, such as an author with writer's block who went on to win literary prizes, and a scientist who will probably be a Nobel Prize recipient. Recognition from one's colleagues is important to almost everybody, and so I am pleased with the honors and awards I have received from various professional groups such as the American Psychological Association, the American Board of Professional Psychology, and the Association for Advancement of Behavior Therapy. I am most proud to have been the recipient of The First Annual Cummings PSYCHE Award for my contributions to the field of behavioral health. Sponsored by the Nicholas & Dorothy Cummings Foundation, this prestigious award included a $50,000 tax-free prize, and a beautifully sculptured bronze statuette of the Greek Goddess PSYCHE. I think of it as "The Oscar of Psychotherapy."

NAJP: What really brings about psychotherapeutic change?

AL: Essentially, three interrelated factors play a role in psychological disturbance – misinformation, missing information, and biological elements. "Misinformation" includes such features as irrational ideas and fallacious (toxic) beliefs. "Missing information" includes various skill deficits, lack of awareness, and especially interpersonal lacunae. At base, we are biochemical/neurophysiological entities, so the biological substratum impacts on most processes. People who have no significant biological imbalances or significant medical concerns (i.e., who are physically healthy) and who also enjoy the benefits of a loving interpersonal network (family, good friends, and pleasant work associates) will have little (if any) need for psychotherapy. Those less fortunate, perhaps with medical illnesses and chemical imbalances, who embrace faulty thinking and have not learned adaptive social and vocational skills, present the main challenges to psychotherapists and counselors. But insight alone, even when coupled with accurate empathy, warmth, and acceptance from a well-intentioned therapist, will do very little to remedy many of the irrational ideas and the deficits that so many clients display. They require coaching, training, rehearsing, and even didactic learning. In short, therapists who dwell on oral, verbal and purely cognitive methods (talk therapy) are far less likely to obtain robust results when compared to those who also focus on behavior change. Indeed, there are data to show that performance-based methods will usually win out over vocal ones. In my treatment, I always ask, "What have you done differently this week?" The old cliché, "actions speak louder than words" is a guiding principle. If a client says he or she has done very little that could be called "different," chances are that there will be no significant progress. I am very happy when a client answers my question with statements such as, "This week I did three things that I have been putting off. First, I told my boss that I am unable to work more than one Saturday a month, and she took it well. I also visited my sister and we settled our argument over the Christmas debacle I had shared with you, and I feel good about this. And thirdly, I kept away from junk food the entire week." My simple aphorism is that (apart from or in addition to whatever biological interventions may be necessary) *real change comes when our clients do things differently and do different things.*

NAJP: How has your approach changed over the years? Or have you remained very eclectic?

AL: I have chapters in two recent books that fully explain the changes that have taken place (*Odysseys in Psychotherapy,* edited by J.J. Shay & J. Wheelis, and *How Therapists Change*, edited by M.R. Goldfried). In essence, in the 1950s I shed most of the shackles of my formal training (mainly psychodynamic, person centered, and Sullivanian approaches) and became enamored of behavior therapy. When follow-ups revealed that my outcomes were less stable or durable than I desired, I became technically eclectic and borrowed promising techniques (not theories) from several disciplines. Thus, in the 1960s, I added elements of Ellis' rational-emotive methods, several Gestalt therapy techniques, and strategies derived from various family therapists, as well as Milton Erickson, Eric Berne, and others. This was not a hodge-podge of subjective methods because I never used any procedure that could not easily be explained by social and cognitive learning theory.

In recent years, my eclecticism has diminished, as I have become reverent towards empirically supported methods – most of which are drawn from the area of cognitive-behavior therapy. I now find that my Multimodal template, the theoretical base of social and cognitive learning theory, and an awareness of biological factors suffice to ensure that my clients' problems are thoroughly assessed and treated with appropriate measures.

NAJP: Do you see audio and video taping clients as important?

AL: In many cases, my answer is affirmative. I have frequently recorded sessions (with the clients' full knowledge and consent, of course). A fair number of clients have brought their own recorders and asked if I would mind if they recorded our sessions. I have never declined. Some of my colleagues have a somewhat paranoid proclivity and are afraid that clients may use the recordings against them. In my experience, it has almost always been valuable when clients went home and listened to a tape of the session. Most of them reported that ideas, suggestions, and issues that had been discussed in the therapy had by-passed them at the time, or they had forgotten them, whereas by listening (sometimes more than once) to the taped session, it refreshed their memory and helped them derive many benefits. When I have listened to tapes of difficult clients, I too have benefited by hearing something that had eluded me, or I have managed to zero in on errors I had made which I would then endeavor to remedy at the next session.

Audio or video recordings are usually extremely useful during behavior rehearsal or role playing procedures. When clients hear and/or see how they come across, it tends to enhance the transfer to the real life situation and ensures that the client is likely to come across more effectively.

NAJP: When would you use group therapy as opposed to individual therapy and why?

AL: Group therapy is often a timesaving and cost-effective approach, especially when one has a relatively homogenous group in search of help for the same problem. For example, a group of people interested in losing weight, or wanting to quit smoking, or seeking to avoid substance abuse may benefit from the peer pressure and the support afforded by a group. I have sometimes found group therapy more effective than individual counseling for people in need of social skills training. Indeed, a man who is afraid of asking women out on dates is likely to obtain a more veridical learning experience when practicing with several woman in a group than when role playing with a male therapist. When working with couples, I have also found it useful to have three or four couples in a group. They seemed to benefit from observing others who were often making the same mistakes as themselves, and the feedback from these other couples (not only the therapist) enhanced the learning process.

NAJP: How important do you see the therapeutic relationship? What are some of the subtle nuances of the relationship?

AL: In some cases, the therapeutic relationship (in the words of the late Carl Rogers) is necessary and sufficient. There are clients who require little more than a good listener, a touch of empathy, or a proverbial shoulder to cry on. In the majority of cases, however, the relationship is necessary but insufficient. Within the context of a positive working alliance, replete with rapport, and genuine respect, most clients require the implementation of treatments of choice. In the heyday of my behavioristic zeal (Johannesburg, circa 1959-1963), I remember seeing several clients who had been under the care of a psychologist (let's call him Dr. G) whom they all regarded as a very fine person. Despite what the psychoanalysts would call an extremely "positive transference," many of his clients showed no significant gains. They derived clear-cut benefits when I administered techniques such as deep relaxation exercises, systematic desensitization, assertiveness training, role-playing, and in vivo assignments. These techniques, of course, were not applied in a disembodied fashion, but first called for the development of adequate rapport. Nevertheless, I would hazard to say that these clients were probably far more attached to and fond of Dr. G than they were to me. In short, a number of writers have stressed that an effective therapist must be a lot more than a nice guy!

NAJP: To what do you attribute most of your therapeutic successes?

AL: This question dovetails nicely with the preceding one. Some years ago an Israeli psychologist interviewed me. He said he had noticed that when effective psychotherapy and successful psychotherapists were topics of conversation, my name was mentioned rather frequently. "Do you consider yourself smarter than most of your peers?" he asked. "No," I answered. "Are you better trained than most of your colleagues?" Again I answered "No." "Well, you must have something special that puts you outside the run of the mill," he insisted. I said I saw the main difference between many of my colleagues and myself as a willingness to go the extra mile. When I first started out as a therapist, I used to go to extremes (e.g., providing free board and lodging in my home for some clients), but as I grew more aware of the need to maintain certain boundaries, I became more circumspect. Nevertheless, compared to most of the professionals I have met, I have remained willing to do more for and with my clients. For example, I exchange emails with clients between sessions (for no extra fee). I encourage my clients to let me know if any questions or issues arise between sessions so that I can think them over beforehand. I am no clock-watcher, and often my sessions will run well over an hour (again at no extra cost to the client). I engage in out-of-the-office experiences with clients, be it dining with an anorexic client to encourage eating, escorting a claustrophobic client to a crowded shopping mall, or attending clients' barmitzvahs, graduations, or weddings when I know that my presence is important for that individual. I believe these niceties fortify the positive aspects of the therapeutic relationship, and potentiate a sense of cooperation and adherence to the therapeutic quest.

REFERENCES

Goldfried, M.R. (Ed).,2001) *How therapists change.* Washington, D.C.: American Psychological Association.

Lazarus, A.A. (1971). *Behavior therapy and beyond.* New York : McGraw-Hill.

Lazarus, A.A. (1976). *Multimodal behavior therapy.* New York : Springer.

Lazarus, A.A. (1985). *Marital myths.* Atascadero, CA: Impact.

Lazarus, A.A. (1989). *The practice of Multimodal Therapy.* Baltimore: The Johns Hopkins University Press.

Lazarus, A.A. (1997) *.Brief but comprehensive psychotherapy: The multimodal way.* New York : Springer.

Lazarus, A.A. (2001). *Marital myths revisited.* Atascadero, CA: Impact.

Lazarus, A.A. & Zur, O. (Eds.) ., (2002). *Dual relationships and psychotherapy.* New York : Springer

Shay, J.J., & Wheelis. (Eds.), (2000). *Odysseys in psychotherapy.* New York: Ardent Media.

Depression, Drug Use, and Gender Differences Among Students at a Religious University

Herbert W. Helm, Jr., Mark D. Boward, Duane C. McBride,
& Rachael I. Del Rio

Andrews University, Berrien Springs

Prior research has indicated a high co-morbidity between depression and substance abuse. The authors investigated the prevalence and relationship of drug use and depression, and how gender affects this relationship, in a conservative religious sample. Overall, this sample had a much lower level of drug use than those reported for national college age populations. Compared to other samples, lower levels of depression were found, and there were no gender differences on the total depression score. Higher levels of depression could account for 17% of the variance of drug use. When Beck Depression Inventory scores were indicative of depression, we could account for almost 5% of the variance within drug use for our female subjects, and 25% of the variance for males. Religiosity was found to have the highest correlation, an inverse one, with drug use.

Depression is the most prevalent mental disorder in the United States (Bromberger & Costello, 1992; NIMH, 2000; Oswalt & Finkelberg, 1995), affecting more than 19 million American adults (National Institute of Mental Health [NIMH], 2000). As noted by Bromberger & Costello (1992), at least one in 20 adults is depressed enough at some point in their life to meet the diagnostic criteria of Major Depressive Disorder, as defined in the *Diagnostic and Statistical Manual of Mental Disorders*, Third Edition–Revised. Not only is depression prevalent among adults, but research suggests that among youth, the prevalence of depression is increasing steadily (Bromberger & Costello, 1992; NIMH, 2000; Oswalt & Finkelberg, 1995; Rosenthal & Schreiner, 2000).

Research indicates there is high co-morbidity (disorders existing simultaneously) between depression and substance abuse (McBride, Van Buren, Terry, & Goldstein, 2000; Ratliff & Burkhart, 1984). It is suggested that, "the rates of depression in drug-using populations exceed those in the general population" (McBride et al., 2000, p. 71). The co-morbidity between depression and substance use disorders is usually attributed to a causal

---

Author info: Correspondence should be sent to: Herbert Helm or Duane McBride, Dept. of Behavioral Science, Andrews University, Berrien Springs, MI 49104.

North American Journal of Psychology, 2002, Vol. 4, No. 2, 183-198.
© NAJP

relationship or an etiological factor shared by both disorders. For example, Swendsen and Merikangas (2000) report that "the association of alcoholism with depression is likely to be attributable to causal factors rather than a shared etiology, but the scarcity of information for other classes of substance use disorders precludes similar conclusions regarding their association with depression" (p. 173).

A review of clinical, epidemiologic, and genetic epidemiologic studies by Swenden & Merikangas (2000) concluded that there was a complex relationship between depression and substance abuse disorders that should be interpreted cautiously. Alcohol was the substance about which they felt a firm conclusion could be drawn. While no single pattern existed, they concluded that for both clinical and community samples, the presence of one created a high risk of the other and that depression and alcohol appear to causally interact. They also found a strong association between depression and other substance use disorders, both licit and illicit. However, due to the types of studies and large number of drug categories, they did not believe there were enough data to completely test specific models of co-morbidity.

A five-year study by Rao, Daley & Hammen (2000) looked at the reciprocal relationship between depression and substance use disorders in women as they made the transition between high school and early adulthood. Not only was there a strong continuity of substance use disorder during this time period, but persons with substance use disorders were at an increased risk of depression. However, the reverse was not true, in that those with depression did not increase their risk of substance use disorders. This suggests that increases in substance use disorders increase the co-morbidity of depression for women.

Over the past thirty years, a 2:1 female-male ratio for depression has been exhibited (Bromberger & Costello, 1992; Culbertson, 1997; Nolen-Hoeksema, Larson & Grayson, 1999; Weissman, Bruce, Leaf, Florio, & Holzer, 1992). Some researchers have reported a ratio of 3:1 female-male ratio for depression (Weissman & Klerman, 1977; Wetzel, 1994). Although most studies consistently report significant differences between genders, Steer and Clark (1997) found that gender variations in depression among a college population were not significantly different for their sample of 107 women [M=11.36, SD=7.07] and 53 men, M=12.85, SD=9.75. Similar studies found that, 18 to 24 year old college men and women exhibit an absence of gender difference in depression (Bryson & Pilon, 1984; Gladstone & Koenig, 1994; Hammen & Padesky, 1977).

As society shifts toward equal gender status, a change in depression scores is notable. It has been suggested that college women may somehow be protected from depressive symptoms (Baron & Matsuyama, 1988; Nolen-Hoeksema, 1990). Baron and Matsuyama (1988) proposed that depression scores become more equal in college environments due to an outgrowth of more equal social roles. Other studies, however, suggest that depression may actually be on the rise among men within college populations due to the effects of gender role conflict (Blazina & Watkins, 1996; Good & Mintz, 1990; Sharpe & Heppner, 1991). Either theory could help explain an absence of gender difference in depression.

Literature suggests that religion serves as a protective factor against drug use. Francis (1992), comparing various Christian denominations, found that affiliation with any Christian religion, and frequent church attendance was correlated with a negative attitude toward alcohol. A negative correlation has been found between religiosity and alcohol use, even after controlling for social class and denomination (Clark, Beeghley, & Cochran, 1990; Mullen, Williams, & Hunt, 1996). Khavari and Harmon (1982) reported that individuals who did not consider themselves religious were more likely to use alcohol, tobacco, marijuana, and amphetamines.

Jessor and Jessor (1977) found that people affiliated with a religion that encourages abstinence from alcohol are more likely to abstain from alcohol use than those affiliated with religions that have a more liberal belief on alcohol consumption. Within religious orientation, researchers have found that fundamental Protestants are less likely to consume alcohol than liberal Protestants, Roman Catholics, (Schlegel & Sandborn, 1979), Lutherans, Jews, or persons with no religion (Boch, Cochran, & Beegley, 1987). Dudley, Mutch, and Cruise (1987) studied a group of Seventh-day Adventist youth and their use of drugs. Seventh-day Adventists are strongly socialized toward abstinence in drugs and alcohol. Compared to a similar aged group they reported a considerably lower frequency of alcohol, tobacco, and marijuana use. A smaller difference was noted with cocaine and amphetamine use.

In a college sample it was found that students with no religious affiliation had higher levels of alcohol use, including frequency of getting drunk, than either Catholics or Protestants. However, orthogonal contrasts indicate that Catholic males show similar drinking behaviors as males in the no religious affiliation group. For Catholics and Protestants extrinsic religiosity seems to play less of a role than intrinsic religiosity in drinking behavior. For Catholics intrinsic religiosity was associated with drinking problems, while for Protestants it was negatively associated with drinking quantity. It appears that social norms and religious affiliation are predictive of alcohol use. However, it may be personality aspects such as high ego involvement (intrinsic religiosity) with a particular religion that may be more predictive of alcohol abuse (Patock-Peckham, Hutchinson, Cheong, & Nagoshi, 1998).

Andrews University is a private Seventh-day Adventist university. Adventism argues that the consumption of any alcohol, tobacco, or other non-medicinal drug is harmful. This belief in abstinence is relayed from generation to generation. Therefore, alcohol, tobacco, and other drug use rates should be low, as these standards of living are advocated in Seventh-day Adventist doctrine. As expected, reports of alcoholism and drug use are somewhat lower among Adventists when compared to mean population rates. While still lower, the frequencies of drugs and alcohol are increasing among Adventists (Hopkins, Hopp, Hopp, Neish, & Rhoads, 1996).

Andrews University has conducted campus-wide drug assessments in 1989, 1991, 1992, 1995, and in 2000. The first three were associated with grants from the Fund for Improvement of Post-Secondary Education (FIPSE). The questions regarding drug prevalence in the 2000 survey closely resembled the wording and content of earlier surveys. It also expanded the contents of the survey to include questions concerning depression, religiosity, locus of control, and the Theory of Reasoned Action. This paper will focus on the depression findings, especially as they relate to drug use and religiosity. We will also compare our drug data to that of the Monitoring the Future Study done at the University of Michigan. Monitoring the Future is a research study that among other topics, studies the prevalence of drug use among college students.

The research questions addressed in this study include the following:

What is the prevalence (trends in drug use) of legal and illegal drug use within this sample, and how does that compare to data from the Monitoring the Future study?

What is the prevalence of depression within this sample, and how does this compare to similar samples of studies using the Beck Depression Inventory?

Do gender differences exist in depression rates?

What is the relationship of drug use and depression, and how does gender affect this relationship?

Does religion seem to be a modifying factor to drug use?

METHOD

Sample

The 2000 survey consisted of 623 subjects. There were 307 females (49.3%) and 312 males (50.1%), of whom the vast majority were single (82.3%). Class standings included 144 freshmen (23.1%), 112 sophomores (18.0%), 128 juniors (20.5%), 134 seniors (21.5%), and 93 graduates/professionals (14.9%). Ethnically, the sample consisted of 312 Caucasians (50.1%), 130 African Americans (20.9%), 88 Asian/Pacific Islanders (14.1%), 45 Hispanics (7.2%), 4 American Indian/Alaskan Natives (0.7%), and 34 Others (5.5%). The majority of the sample was between the ages of 20-23 (48.1%), approximately 25% were below the age of 20, and 26% were at least 24 years old. There were 347 on-campus residential students (55.7%) and 273 off-campus or community students (43.8%). Approximately 65% had attended the university for two years or less. Five Hundred and Three (80.7%) were raised in a family structure where both parents were in the same house. The vast majority, 90.7%, reported the same religious affiliation as the denominationally-run university. Where the numbers for a category do not add up to 623, or the percent does not equal 100, it is usually due to missing values.

Procedure

It was decided to use certain time periods during the school day to collect the data. These time periods were chosen based on the large number of classes being offered and the large number of students likely to be taking those classes. Professors teaching a class during one of those periods were contacted for permission for the survey to be given during one of their class periods. If the professor agreed, a group of trained survey proctors, were sent to give the survey during one of four days. A letter that described the terms, procedure, and information about the survey was read verbatim to each class, as well as a copy put into each survey packet that was passed out. The letter explained that participation was voluntary and that data gathered from the questionnaire was completely confidential. This study had IRB approval.

Materials

A questionnaire was developed which consisted of 123 questions. As noted earlier, this version of the questionnaire closely resembled the wording and content of earlier surveys. The questions asked about demographics and issues such as: the rates of legal and illegal drug use, alcohol consumption, unsafe sexual practice, the attitude of individuals toward these practices, as well as, attitude and perceptions of their immediate environment. In addition to the questions revolving around drug use and related behavior, the questionnaire assessed four areas: depression, religiosity, locus of control and the Theory of Reasoned Action. In this paper, locus of control and the Theory of Reasoned Action will not be addressed.

Depression was measured using the *Beck Depression Inventory* (BDI) (Beck, Ward, Mendelson, Mock, & Erbaugh, 1961). The BDI consists of 21 questions. "Thirteen items cover cognitive and affective components of depression such as pessimism, guilt, crying, indecision, and self-accusations; eight items assess somatic and performance variables such as sleep problems, body image, work difficulties, and loss of interest in sex" (Gregory, 1996, p. 562). Each question has four possible answers of which each is assigned a value from 0 to 3. All points or values are added together to get a composite score. The range of possible scores is 0 - 63. According to the manual (Beck & Steer, 1993), the Center for Cognitive Therapy is recommending the following cutoff scores: 0-9 for the minimal range; 10-16 for the mild range; 17-29 for the moderate range; and 30-63 for the severe range.

Based on a study by Oliver and Simmons (1984), the scoring guidelines in the BDI manual suggests that scores greater than 15 can be indicative of possible depression. They also recommend that this be confirmed by a trained clinician. Revised BDI scores do not seem to be meaningfully related to gender or age. A twenty-five year evaluation of the inventory's psychometric properties and demographic correlates can be found in Beck, Steer, and Garbin (1988).

A religiosity scale was developed using eleven questions that asked about participants' level of participation in various religious activities. These activities were placed on a four-point Likert scale that ran from "never" to "regularly." They included behaviors such as: personal prayer, attending church and Sabbath school services, reading Biblical and other religious writings, family worship at home, telling someone about your faith, attending religious clubs, and attending church-sponsored social events.

A new variable was computed from the survey to assess drug use within the past year. It was developed by combining the reported number of drugs surveyed, with the frequency that subjects used a given substance. As expected, subjects would get a higher value on this scale as the frequency with which they used a drug and/or the number of drugs used increased. Unless noted otherwise, this is the variable known as the "drug use" variable.

RESULTS

Prevalence of Legal and Illegal Drugs Use

Table 1 compares the trends of various types of drugs, at thirty-day and annual time periods. The comparison is between the findings from the 2000 Andrews University survey and those of the 1999 Monitoring the Future Study (Johnston, O'Malley, & Bachman, 2000). The Monitoring the Future study was conducted at the University of Michigan's Institute for Social Research. We are comparing our data to that of their college sample. Their college sample consisted of full-time students at two- or four-year colleges, who are one to four years beyond high school. The approximate weighted *N* is equal to 1440.

As can be seen from Table 1, the sample at Andrews University has a much

TABLE 1 Trends in the Prevalence of Various Types of Drugs

	AU '00 last 30 days	MtF '99 last 30 days	% difference	AU '00 last year	MtF '99 last year	% difference
Tobacco	10.8	30.6	-19.8	16.0	44.5	-28.5
Alcohol	20.2	69.6	-49.4	34.5	83.6	-49.1
Marijuana	5.5	20.7	-15.2	11.1	35.2	-24.1
Cocaine	0.6	1.2	-0.6	1.5	4.6	-3.1
Amphetamines	2.6	2.3	0.3	4.3	5.8	-1.5

MtF = Monitoring the Future
(Entries are percentages)

lower level of drug use, on the five drugs sampled, for each of the time periods. The exception to this is amphetamine use for the last thirty days. It should be noted that the Monitoring the Future study reports an "adjusted" amphetamines percentage. This meant that the drug use was not reported if under a doctor's order, and the data was from a revised question "which attempts to exclude the inappropriate reporting of non-prescription stimulants "(p. 217). The Andrews survey simply said: "Amphetamines (diet pills, speed)." Differences may have occurred due to survey methodology on this question.

Alcohol and tobacco continue to be shown as the most frequently used drugs at Andrews University, with alcohol being the most frequently used drug within the past year. The most frequently noted age of first use was between 17-20 years old for alcohol, marijuana, cocaine, and amphetamines. For tobacco, the most frequently noted age of first use was between 13-15 years old.

Depression

There were 623 subjects who completely filled out the *Beck Depression Inventory*. Our subjects scored as follows on the *Beck Depression Inventory*: 487 had scores within the minimal range (0 -9), or 78.2%; 81 subjects had scores within the mild range (10-16), or 13.0%; 50 subjects had scores within the moderate range (17-29), or 8.0%; and 5 subjects had scores within the severe range (30-63), or 0.8% of our sample. The mean score for our sample was 5.94 and the standard deviation was 6.73. Based on the guidelines provided by Beck and Steer (1993), our sample mean would not indicate symptoms of depression.

In Table 2, the mean and standard deviation of this sample are individually compared to three other recent BDI samples (Arthur & Hayward, 1997; Endler, Rutherford, & Denisoff, 1999; Killgore, 1999). *T* tests were conducted between the Andrews University sample and each of the other three samples noted above. As can be seen on Table 2, there is significantly less depression with the Andrews University sample. The study by Arthur and Hayward (1997) found that 32% of their sample reported depression levels in the mild to severe range. The Andrews University sample had 21.8% that reported mild to severe levels of depression, as measured on the BDI.

TABLE 2 A Comparison of Three Other Samples and Levels of Depression on the BDI

Sample:	Mean	SD	*t*
Killgore, 1999; N = 303	8.41	7.28	5.51**
Arthur & Hayward, 1997; N = 178	7.61	6.60	2.96*
Endler et al., 1999; N = 565	9.30	7.32	8.21**
Andrews Univ., 2000; N = 623	5.94	6.73	

* $p < .01$. ** $p < .001$.

Gender Differences With Depression

Five of the subjects with completed BDIs were missing gender identification, therefore, they were excluded and $N = 619$. The mean score for the 312 men in our survey was 5.50 (SD = 7.13), and for the 307 females, the mean was 6.38 (SD = 6.30). A significant difference was not found between these two scores [$t_{(617)} = -1.64, p = .101$]. In comparison, a re-analysis of the study by Arthur and Hayward (1997) shows a significant difference between genders on the BDI [$t_{(565)} = 3.04, p < .01$]. Their 400 female students had a mean of 9.90 (SD = 7.29), and their 165 male students had a mean of 7.87 (SD = 7.19), a difference of slightly more than two points. As previously mentioned, Table 2 indicates lower overall depression scores when compared to three other samples, this holds up with gender differences. When comparing our sample to that of Arthur and Hayward (1997) both males [$t_{(475)} = 3.43, p < .001$] and females [$t_{(705)} = 6.99, p < .001$] were significantly lower in our sample.

While no significant difference was found between the total depression score on the BDI for gender, there were some significant differences when each depression category was individually analyzed for gender differences. There was a significant difference when looking at the minimal category, scores of 0-9, [$t_{(483)} = -3.4, p = 001$]. The 252 males had a mean score of 2.62 (SD = 2.63) and the 233 females had a mean score of 3.46 (SD = 2.81). No significant gender differences were found within the mild range, 10-16 [$t_{(77)} = .27$, ns]. The 33 males had a mean score of 12.33 (SD = 2.09) and the 46 females had a mean score of 12.21 (SD = 1.71). A gender difference is also noted in the moderate range [$t_{(44)} = 2.29, p = .027$]. The numbers given are for equal variances not assumed, as Levene's test for equality of variances was significant, [$p = .006$]. The 24 males had a mean of 22.38 (SD = 3.76) and the 26 females had a mean of 20.15 (SD = 3.02). Since there were only five subjects (three males and two females) at the severe level of depression, it would not be worthwhile to run statistics. While two of the categories do show gender differences, these are not likely to affect most of the tests we run since we are either running them against the total number of complete BDI's or those indicative of depression (>15). There is no significant gender difference when comparing those having scores indicative of depression, [$t_{(59)} = 1.48$, ns]. The mean score for the 31 males was 22.94 (SD = 6.63) and the mean score for the 30 females was 20.77 (SD = 4.59).

Drug Use and Depression

The relationship between the drug use variable and the BDI, reveals a Pearson's correlation of .23 ($p < .001$, $N = 615$). This accounts for approximately 5% of the variance. When looking at total depression scores of greater than 15, those scores which may be indicative of depression, the

Pearson's correlation is .41 ($p = .001$, $N = 61$). The increase in correlation with higher levels of depression means that we can account for more of the variance of drug use. Here we account for approximately 17% of the variance. This suggests that as persons score higher on the BDI, the probability exists of more frequent drug use.

When the relationship between drug use and depression was evaluated based on gender differences, an even more interesting picture developed. When using scores for males which may be indicative of depression (>15), a Pearson's correlation of .497 was found ($p = .004$, $N = 31$). The mean depression score for this group was 22.94 ($SD = 6.63$). This meant that we could account for almost 25% of the variance within drug use for our male subjects who had BDI scores indicative of depression. When females in our study were analyzed using the same criteria, a Pearson's correlation of .218 was found (p was not significant, $N = 30$). The mean depression score for females was 20.77 ($SD = 4.59$). Here we could account for almost 5% of the variance within drug use for our female subjects with BDI scores indicative of depression. This would indicate that depression plays a more essential role in the frequency of drug use for males in this study, than for females, who show scores indicative of depression.

Chi-square analysis was done between depression categories and whether or not the subjects had reported drug use in the last year. Depression scores were re-coded into one of three categories: minimal (scores of 0-9), mild (scores of 10-16), and moderate/severe (scores of 17 and greater). The moderate and severe categories were collapsed due to the small number of subjects in the severe category of depression. Due to the low reported use of various drugs reported, drug categories were collapsed into either a "no use" category or an "any use" category, within the last year. Statistical significance ($p \leq .001$) with depression was found for tobacco ($X^2 = 22.55$), alcohol ($X^2 = 16.24$), and marijuana ($X^2 = 13.08$). A consistent pattern was found for all three drugs with depression. As depression scores went up there was a greater than expected occurrence (increase) of reported drug use, while at the lowest level there was less reported use than would be expected, however, a causal relationship is not being discussed. When comparing whether or not alcohol use had occurred in the past year and levels of depression, the most interesting finding was that there was no statistical significance for females when analyzed by gender. Due to the reported low use of cocaine and amphetamines, analysis was not done, due to low expected counts for some cells.

Drug Use and Religiosity

As noted earlier a number of studies suggest that religion seems to be a protective factor to drug use. In this study, a forward multiple regression model was used with religiosity, depression, and gender as the independent variables and drug use as the dependent variable. Table 3, Correlations With Drug Use, reports correlations among the independent variables, and with the dependent variable. Religiosity has the highest correlation, an inverse one with drug use.

TABLE 3 Correlations With Drug Use

	1	2	3	4
1. Drug Use				
2. Religiosity	-.41*			
3. Depression	.23*	-.25*		
4. Gender	-.07	-.03	.07	

*p <.01

By itself religiosity accounted for 17.2% of the variance in drug use. When depression and gender are added, another 2.5% of the variance could be accounted for, or 19.7%. When religiosity was removed from the model and the other three variables were used; we could account for only 6.5% of the variance. Based on the predictors that were used, religiosity was the best predictor of reduced drug use.

DISCUSSION

The purpose of this study was to examine the relationship between depression and substance use at a conservative church-affiliated university that explicitly forbids the use of alcohol, tobacco or other drugs by students. The depression and substance use relationship was also examined by gender. The study and the analysis reported in this paper is within the context of a wide variety of recent research, as noted earlier, that has been done on the relationship between depression, religiosity, gender and drug use. Many researchers have documented that depression is one of the most frequently occurring and undiagnosed health problems in the United States (NIMH 2000). For decades, researchers have also suggested that depression may have significant co-morbidity with substance use (McBride et. al., 2000). Traditionally, researchers and clinicians have reported that females were significantly more likely to have higher rates of depression. However, as was noted earlier, gender differences in depression may be changing as a result of major changes in gender social roles. While research on the relationship between depression, gender and drug use has often been undertaken in general society, it has rarely been done in a subculture that prohibits substance use. Examining these issues in a conservative church affiliated university population provides the opportunity to examine the existence and strength of a relationship between depression, drug use and gender within the context of a prohibitionist subculture.

While the reported use rates of alcohol, tobacco and other drugs reported in this study are generally two-thirds less than those reported for national college age populations, the analysis showed that even in this prohibitionist subculture there was substance use. About one-third of the students reported alcohol use in the last year, about 16 percent reported tobacco use, and 11 percent marijuana use. These data may suggest the pervasiveness of mass culture. Any group or subculture likely finds it very difficult to isolate itself from general cultural patterns. Cultural leveling tends to occur between subcultures and mass culture. Social scientists have argued that cultural leveling – the virtual disappearance of attitude and behavioral differences between social groups – is a characteristic of a post-modern age (Kubiak, 1999). Certainly one of the main issues in any distinctive subculture is the maintenance of the differences between itself and the larger culture. The substance use data reported in this paper suggest that Adventist culture has had a measure of success in maintaining a distinctive difference in drug use patterns in comparison to general cultural behavior. This difference is noted both within this study and the study by Dudley, Mutch, and Cruise (1987). However, the data also suggest some cultural leveling in drug use patterns. Cultural leveling suggests that subcultures are affected by the general cultural around them.

The depression scores found in the study sample appear to be low. A comparison to other college populations reported in the literature indicates that students at this university had lower depression scores than students at other colleges. Understanding the reasons for these differences would require data beyond that which is available in this study. Researchers have argued that religion has a protective association with depression (Drumm, 2002). That is, those who are more involved in religion have lower rates of depression. Generally students in this study had high religious involvement and there was a correlation of -.41 between religiosity and depression. It may be that religiosity rates are so high in this population that depression scores are lower as a result.

While both substance use rates and depression scores appear to be lower in this Adventist population, analysis found a moderate relationship between depression and substance use within the last year. About five percent of the variance in substance use in the last year was explained by subjects depression scores. The analysis further showed that for those with higher depression scores, there was a much stronger relationship between depression and substance use. For those who had a BDI score of 15 or higher, about 17 percent of the variance in substance use was explained by depression scores. Overall these data suggest that even in this subcultural context of less substance use, lower depression rates and high rates of religiosity, depression is still an important variable in understanding the use of alcohol, tobacco and other drugs.

Perhaps the most important finding in this analysis was the differences in the relationship between depression and substance use by gender. Traditionally, researchers have found a higher rate of depression among females and a stronger relationship between depression and drug use for females (McBride et al., 2000). These data show no differences in total depression scores by gender and that depression explains a much higher proportion of alcohol use among males than among females. This finding is difficult to interpret and these data do not lend to sorting out the time sequence behavior that might best address this finding. It may be that religious involvement has more of a depression protective effect for females than for males. It may also be that in a conservative subculture males are more likely to engage in higher risk behavior (drug use) than females. Regardless of the reasons for these findings, the data suggest some important questions that need to be addressed. These include examining depression scores in other conservative religious subcultures, trends in depression scores by gender among young people in general and, perhaps, particularly among conservative religious populations, and the differences in the protective effect of religion by gender. With the current emphasis on the role of religious belief in understanding and ameliorating a wide variety of social problems (Pattillo-McCoy, 1998; Shank & Reed, 1995), including drug use, it is important to examine these questions further. This would include further study into which factors within religious belief have the most protective factors relating to drug use.

REFERENCES

Arthur, N., & Hayward, L. (1997). The relationships between perfectionism, standards for academic achievement, and emotional distress in postsecondary students. *Journal of College Student Development, 38*(6), 622-632.

Baron, M., & Matsuyama, Y. (1988). Symptoms of depression and psychological distress in United States and Japanese college students. *Journal of Social Psychology, 128*, 803-816.

Beck A. T., & Steer, R. A.(1993). *Beck depression inventory manual*. San Antonio, TX: Psychological Corporation, Harcourt, Brace, & Co.

Beck, A. T., Steer, R. A., & Garbin, M. G. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review, 8*, 77-100.

Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry, 4*, 561-571.

Blazina, C., & Watkins, C. E., Jr. (1996). Masculine gender role conflict: Effects of college men's psychological well-being, chemical substance usage, and attitudes towards help-seeking. *Journal of Counseling Psychology, 43*(4), 461-466.

Bock, E., Cochran, J., & Beeghley, L. (1987). Moral messages: The relative influence of denomination on the religiosity-alcohol relationship. *Sociological Quarterly, 28*, 86-105.

Bromberger, J., & Costello, E. (1992). Epidemiology of depression for clinicians. *Social Work, 37*(2) 120-126.

Bryson, S., & Pilon, D. (1984). Sex differences in depression and the method of administering the Beck Depression Inventory. *Journal of Clinical Psychology, 40*(2), 529-534.

Clark, L., Beeghley, L., & Cochran, J. (1990). Religiosity, social class and alcohol use: An application of reference groups theory. *Sociological Perspectives, 33*, 201-218.

Culbertson, F. (1997). Depression and gender: An international review. *American Psychologist, 52*(1), 25-31.

Drumm, R. (2002). Protective effects of religion: Drug use, HIV risk, and violence. Research in support of current health policy initiatives. *Journal of Addictions Nursing, (*in press*).*

Dudley, R., Mutch, P., & Cruise, R. (1987). Religion factors and drug use among Seventh-day Adventist youth in North America. *Journal for the Scientific Study of Religion, 26*, 218-233.

Endler, N. S., Rutherford, A., & Denisoff, E. (1999). Beck Depression Inventory: Exploring its dimensionality in a nonclinical population. *Journal of Clinical Psychology*, 55(10), 1307-1312.

Francis, L. L. (1992). Attitude toward alcohol, church attendance, and denominational identity. *Drug Alcohol Dependence, 31*, 45-50.

Gladstone, T. R., & Koenig, L. J. (1994). Sex differences in depression across the high school to college transition. *Journal of Youth and Adolescence, 23*(6), 643-669.

Good, G., & Mintz, L. (1990). Toward healthy conceptions of masculinity: Clarifying the issues. *Journal of Mental Health Counseling, 15*, 403-413.

Gregory, R. J. (1996). *Psychological testing, history, principles, and applications* (2^{nd} ed.). Boston, MA: Allyn and Bacon.

Hammen, C., and Padesky. (1977). Sex differences in the expression of depressive responses on the Beck Depression Inventory. *Journal of Abnormal Psychology, 86*, 609-614.

Hopkins, G. L., Hopp, J. W., Hopp, H. P., Neish, C., & Rhoads, G. (1996). AIDS and Adventist youth. *Ministry, 69*, 22-27.

Jessor R., & Jessor S. (1977). *Problem Behaviour and Psychological Development: A Longitudinal Study of Youth.* New York: Academic Press.

Johnston, L. D., O'Malley, P. M., & Bachman, J. G. (2000). *Monitoring the Future National Survey Results on Drug Use, 1975-1999: Vol. 2. College students and adults ages 19-40.* National Institute on Drug Abuse, NIH Publication No. 00-4803. Bethesda, Maryland.

Khavari, K. A., & Harmon, T. M. (1982). The relationship between the degree of professed religious belief and the use of drugs. *International Journal of the Addictions, 17*(5), 849-857.

Killgore, W. S. (1999). Empirically derived factor indices for the Beck Depression Inventory. *Psychological Reports, 84*, 1005-1013.

Kubiak, A. E. (1999). The new age, postmodern conspiracy. *Social Compass, 46*(2), 135-143.

McBride, D. C., Van Buren, H., Terry, Y., M., & Goldstein, B., J. (2000). Depression, drug use and health services need and utilization. *Emergent Issues in the Field of Drug Abuse.* J. Levy, R. Stephens & D. McBride (Eds.), *Advances in Medical Sociology.* Stamford, Connecticut: JAI Press.

Mullen, K., Williams, R., & Hunt, K. (1996). Irish descent, religion, and alcohol and tobacco use. *Addiction, 91*(2), 243-254.

National Institute of Mental Health. (2000). *Depression Research at the National Institute of Mental Health: Fact Sheet.*

Nolen-Hoeksema, S. (1990). Sex differences in depression. *Journal of Health and Social Behavior, 27*, 161-178.

Nolen-Hoeksema, S., Larson, J., & Grayson, C. (1999). Explaining the gender difference in depressive symptoms. *Journal of Personality & Social Psychology, 77*(5), 1061-1072.

Oliver, J. M., & Simmons, M. E. (1984). Depression as measured by the DSM-III and the Beck Depression Inventory in an unselected adult population. *Journal of Consulting and Clinical Psychology, 52*, 892-898.

Oswalt, R., & Finkelberg, S. (1995). College depression: Causes, duration, and coping. *Psychological Reports, 77*, 858.

Pattillo-McCoy, M. (1998). Church culture as a strategy of action in the black community. *American Sociological Review, 63*, 767-784.

Patock-Peckham, J., Hutchinson, G., Cheong, J., & Nagoshi, C. (1998). Effects of religion and religiosity on alcohol use in a college student sample. *Drug and Alcohol Dependence, 49*, 81-88.

Rao, U., Daley, S. E., & Hammen, C. (2000). Relationship between depression and substance use disorders in adolescent women during the transition to adulthood. *Journal of the American Academy of Child & Adolescent Psychiatry, 39(2)*, 215-222.

Ratliff, K. G., & Burkhart, B. R. (1984). Sex differences in motivations for and effects of drinking among college students. *Journal of Studies on Alcohol, 45*(1), 26-32.

Rosenthal, B. S., & Schreiner, A. C. (2000). Prevalence of psychological symptoms among undergraduate students in an ethnically diverse urban public college. *Journal of American College Health, 49*(1), 12-19.

Schlegel, R., & Sandborn, M. (1979). Religious affiliation and adolescent drinking. *Journal of Studies on Alcohol, 40*, 693-703.

Sharpe, M., & Heppner, P. (1991). Gender role, gender-role conflict, and psychological well-being in men. *Journal of Counseling Psychology, 38*, 323-330.

Shank, H., & Reed, W. (1995). A challenge to suburban evangelical churches: Theological perspectives on poverty in America. *Journal of Interdisciplinary Studies, 7*, 119-134.

Steer, R. A., & Clark, D. A. (1997). Psychometric characteristics of the Beck Depression Inventory - II with college students. *Measurement and Evaluation in Counseling & Development, 30*(3), 128-136.

Swendsen, J. D., & Merikangas, K. R. (2000). The comorbidity of depression and substance use disorders. *Clinical Psychology Review, 20*(2), 173-189.

Weissman, M., Bruce, M., Leaf, P., Florio, L., & Holzer, C. (1992). Affective disorders in psychiatric disorders in America: The Epidemiological Catchment Area Study. In L.N. Robins & D. A. Regier (Ed.), *Psychiatric Disorders in America: The Epidemiological Catchment Area Study* (pp. 53-80). New York: Free Press.

Weissman, M. M., & Klerman, G. L. (1977). Sex differences and the epidemiology of depression. *Journal of Health and Social Behavior, 26*, 156-182.

Wetzel, J. W. (1994). Depression: Women at risk. *Social Work–Health Care, 19*, 85-108.

Note: Herbert W. Helm Jr, Ph.D., is an Associate Professor of Psychology at Andrews University.

Mark D. Boward, M.A., is a Research Assistant at Andrews University.

Duane C. McBride, Ph.D., is a Research Professor of Sociology at Andrews University.

Rachael I. Del Rio, is an undergraduate Research Assistant at Andrews University.

Participation Motives in Predicting Sport Dropout Among Australian Youth Gymnasts

Todd A. Ryska & David Hohensee
University of Texas at San Antonio

Dean Cooley and Christine Jones
University of Tasmania

The rate of attrition within Australian competitive youth sport has been an area of continuing concern among sport personnel and researchers. A greater understanding of this issue may be achieved through studying the patterns of participation motives evidenced by these young athletes. Among a sample of Australian gymnasts ($N = 281$), logistic regression analyses indicated that age and the participation motives of fitness, team atmosphere, new situation, social recognition, and self-challenge reliably distinguished athletes who subsequently continued or withdrew from competitive gymnastics. Athletes who eventually dropped out had participated in gymnastics for largely extrinsic reasons, whereas those athletes who continued their participation were more intrinsically involved. Although sport dropout has been largely linked to a lack of intrinsic qualities in the sport setting, it also appears that an emphasis on extrinsic outcomes by some athletes contributes to their subsequent withdrawal from competitive sport. Implications of the present results are discussed in terms of the costs-benefits component of Gould's (1987) youth sport attrition model.

The competitive sport experience occupies a valued place in Australian culture and is considered an especially important aspect of personal development among young people. The positive benefits of competitive sport involvement has the potential of reaching the approximately 1.6 million Australian children and adolescents who participate in organized competitive sport programs (Australian Bureau of Statistics, 2001). Research interest in the youth sport participant has undoubtedly increased as a result of the Olympic Games hosted in Sydney, with much attention centering upon the motives for participation in, and departure from, competitive youth sport. Although the yearly dropout rate in competitive youth sport is estimated at 35% worldwide, significantly higher rates have been reported among Australian children

---

Author info: Correspondence should be sent to: Todd A. Ryska, PhD, University of Texas at San Antonio, College of Education, San Antonio, TX 78249. E-mail: tryska@utsa.edu.

North American Journal of Psychology, 2002, Vol. 4, No. 2, 199-210.

© NAJP

(Clough & Traill, 1992; Robertson, 1992). Particular concern has been raised among athletic and counseling personnel regarding the elevated attrition rates evidenced among young Australian gymnasts (Australian Sports Commission, 1998, 1999). A consistent and positive competitive environment is especially important for these athletes given the comparatively high physical and mental demands placed on them at a relatively young age. Through ascertaining the specific demographic factors and motivational patterns which place these athletes at risk for dropout, athletic personnel may become more empowered to create such an environment.

Initial evidence suggests that athlete sex, age, and competitive experience are influential factors in the sport withdrawal process. Sex differences in motives have been reported with boys being more likely to discontinue their participation given the competitive philosophy of the particular program as well as time conflicts with other sports (Weiss & Glenn, 1992). Athletes of varying ages also tend to report different reasons for withdrawing from sport. Whereas younger sport dropouts typically cite reasons associated with the program structure (e.g., little success, lack of playing time, overemphasis on winning) and the social context of the sport (e.g., no teamwork, did not meet new friends, did not feel important enough), older athletes report conflicting interests between sport and nonsport activities (e.g., no free time, conflicts with social life, disrupts work schedule) (Fry, McClements, & Sefton, 1981; Petlichkoff, 1992). Participation motives appear to vary as a function of competitive level as well (Klint & Weiss, 1986; Weiss & Petlichkoff, 1989). Highly competitive gymnasts tend to report competence, fitness, and challenge as participation motives, whereas recreational gymnasts report that fun and situational motives are most important. Although no work to date has directly assessed the role of athlete sex, gender, and competitive experience on participation motives within Australian competitive gymnastics, related studies generally suggest that the participation motives of increased physical fitness, competitive experience, and skill improvement are reported as influential among Australian youth sport participants (Clough & Traill, 1992; Robertson, 1992).

Descriptive studies of youth sport attrition reveal that although children reveal a variety of reasons for discontinuing sport participation, such motives may be generally categorized as either intrinsic or extrinsic in nature. Intrinsic sport motivation is generally defined as an inherent desire and curiosity of pursuing optimal skill challenges within the sport setting which often result in enhanced personal competence, perceived control, and self-confidence. Numerous motivation theorists suggest that intrinsic reasons for sport attrition center upon the themes of being bored, not having fun, and not improving one's skills (Burton & Martens, 1986;

Gould, Feltz, Horn, & Weiss, 1982; Petlichkoff, 1992). In contrast, extrinsic sport motives are largely determined by external sources typically in the forms of social approval, material rewards, and an emphasis on interpersonal superiority. Extrinsic reasons for sport withdrawal tend to focus on coaching behaviors and the sport program structure such as an overemphasis on winning and inadequate playing time (Athletic Footwear Association, 1990; Klint & Weiss, 1986). Strong evidence indicates that individual differences in sport motive impact a variety of sport-related behaviors. As compared to their extrinsically motivated counterparts, young athletes who participate in sport for primarily intrinsic reasons have been shown to expend greater effort and persist longer (Duda, 1989), make adaptive attributions for success and failure (Hall, 1990), report greater self-confidence (Seifriz, Duda, & Chi, 1992), and properly identify with their sport involvement (Weitzer, 1989). Although considerably less is known regarding the impact of participation motives on attrition within competitive gymnastics, these results might be extended to suggest that intrinsically motivated gymnasts would be less likely to drop out of their sport than gymnasts motivated to participate by largely extrinsic reasons.

Youth sport research has met with difficulty in discerning a pattern in attrition behavior due to various ways the construct has been conceptualized. For example, some studies have defined dropouts as those athletes who have discontinued their participation in the sport under study. This approach does not account for those athletes who move to another sport or, alternately, to a lower competitive level of the same sport. It would be reasonable to suspect that the participation motives underlying these different types of "dropout" may vary (Gould & Medbery, 2000; Petlichkoff, 1993). Given the methodology of previous research, it is understandable that tracking athletes who ultimately drop out of competitive sport completely would be difficult. Unfortunately, it is this population that may be most at-risk of departing from sport in a personally disruptive manner (Burton, 1988; Seefeldt, Ewing, & Walk, 1993).

The present study sought to extend the youth sport attrition research by distinguishing current and former youth gymnasts on the basis of their participation motives above and beyond the effects of sex, age, and competitive experience. In particular, intrinsic motives such as fitness, action, and challenge as well as extrinsic motives including competition and social recognition were assessed to determine their ability to predict athletes' participation status. Current findings in youth sport motivation and attrition research would suggest that athletes whose participation motives are predominately extrinsic in nature would be more likely to quit sport than their more intrinsically involved counterparts. However, due to the exploratory quality of the present investigation, specific hypotheses regarding the predictive ability of individual motives were not forwarded.

METHOD

Sample and Procedure

The sample consisted of 349 youth gymnasts who were currently registered with the Australian Coaching Council (ACC) and participating on a state level gymnastics team. This sample ranged in age from 9 to 18 years (*M* = 14.5, *SD* = 3.2) and reported 5.9 (*SD* = .71) years of competitive gymnastics experience. Each of the participants trained and competed in the areas of vault, bars, beam, and floor routine. Given the ages of the sample, informed consent was obtained from the legal guardian of each athlete and was verified by the principal investigator before permission to participate was granted. A survey packet including a covering letter, questionnaires, and a debriefing form was mailed to the sample. All participants were informed as to the anonymity and confidentiality of their responses, as well as their right to terminate involvement at any time.

Each athlete's participation status for the subsequent competitive season was tracked through the ACC athlete database. Of the original sample, 149 gymnasts ranging in age from 9 to 18 years (*M* = 13.7, *SD* = 3.1) remained registered and active on ACC sponsored gymnastics teams a year later. In contrast, 132 athletes ranging in age from 12 to 18 years (*M* = 16.3, *SD* = 2.7) had withdrawn from ACC sponsored competitive gymnastics and remained inactive in any sponsored sport for a two year period.

Measures

Demographic data. Athletes were asked to provide various biographical data including sex, age, current and highest competitive levels obtained in gymnastics, and amount of competitive experience.

Participation motives. Athletes' underlying reasons for engaging in competitive gymnastics were measured by the Motives for Participating in Gymnastics Scale (Klint & Weiss, 1987). Athletes were asked to respond to 32 statements regarding their reasons for gymnastics participation along a 5-point scale anchored by 1 (*not at all important*) and 5 (*extremely important*). The reliability and construct validity of the scale have been substantiated across a variety of sports (Gould, Feltz, & Weiss, 1985; Klint & Weiss, 1987).

RESULTS

The means, standard deviations, and reliability estimates for the participation motives for both participants and dropouts are listed in Table 1. Similar to previous findings (Gould et al., 1985; Klint & Weiss, 1986), the seven factors of competition (e.g., like to compete, receive rewards), action (e.g., get rid of energy, like the excitement), fitness (e.g., stay in shape, like to exercise), team atmosphere (e.g., like the teamwork like the team spirit), new situation (e.g., something to do, reason to get out of the house), social recognition (e.g., feel important, others notice me), and challenge (e.g., learn new skills, perform a good skill) accounted for 84.8% of the variance in participation motives.

TABLE 1 Descriptive Statistics for the Measured Variables Among Participants and Dropouts

	Participants			Dropouts		
Variable	M	SD	Alpha	M	SD	Alpha
Competition	3.78	.43	.75	3.89	.39	.77
Action	3.66	.42	.82	3.58	.40	.85
Fitness	4.29***	.51	.77	3.52	.44	.78
Team atmosphere	3.92**	.39	.86	3.21	.41	.84
New situation	2.48	.29	.88	3.54**	.32	.82
Social recognition	3.37	.46	.84	4.12**	.47	.80
Personal challenge	4.32**	.40	.72	3.29	.39	.74

Note. ** p < .01. *** p < .001.

Although it is common practice in sport science research to conduct a series of univariate *F* tests as a follow-up procedure to a significant MANOVA main effect, several authors have cautioned against this statistical approach (Huberty & Morris, 1989; Maxwell, 1992). Marsh (1994) states that under conditions in which variables represent a common construct, the calculation of univariate *F* ratios may be inappropriate in that such ratios would not account for the moderate to high intercorrelations found among multifactor variables. Thus, a sequential logistic regression analysis was conducted to assess the prediction of participation status (i.e., participant or dropout) first on the basis of demographic factors (i.e., sex, age, competitive experience) and then after addition of the seven participation motives. This approach provides direct evidence as to whether the participation motives significantly enhance the prediction of participation status after contribution of the demographic variables. Adequate group discrimination (i.e., model fit) was demonstrated on the basis of the demographic variables only, [$\chi 2$ = 682.33, ns], and after addition of the seven participation motive predictors, [$\chi2$ = 668.02, *ns*]. A comparison of the log-likelihood ratios for models with and without the participation motives indicated reliable improvement in the model with the addition of the participation motives, [$\chi2$ = 35.51, p = .04].

TABLE 2 Logistic Regression Analysis of Participation Status as a Function of Demographic and Participation Motive Variables

Variable	χ^2 to Remove	Sig. p	Log-Likelihood
Demographic measures			
Sex	5.76		
Age	18.23	.002	
Experience	0.95		
All demographic variables			-268.77
Participation motives			
Aspects of competition	1.55		
Action	0.75		
Fitness	17.43	.002	
Team atmosphere	27.44	.001	
New situation	10.29	.004	
Social recognition	8.46	.004	
Challenge	14.86	.003	
All variables			-251.02

On the basis of the three demographic variables, 48% and 72% of the participants and dropouts were correctly predicted, respectively, for an overall correct prediction of 65%. Prediction success improved to 82% with the addition of the seven participation motives for success rates of 55% and 84% for the participant and dropout groups, respectively. Maximum likelihood tests were run for each predictor to assess whether the model would be significantly degraded by removal of each predictor. An alpha criterion level of .005 was set to combat Type I error (Tabachnick & Fidell, 1996). As indicated in Table 2, six predictors reliably distinguished participants from dropouts, including age and the participation motives of fitness, team atmosphere, new situation, social recognition, and personal challenge. Comparison of mean group differences indicated that dropouts tended to be older than participants, and were more motivated to participate in competitive gymnastics for situational (e.g., 'get out of the house') and social recognition reasons (e.g., 'be popular with others'). On the other hand, participants continued their competitive gymnastics for reasons of fitness (e.g., 'stay in shape'), team atmosphere (e.g., 'like the team spirit'), and personal challenge (e.g., 'learn new skills').

DISCUSSION

The participation motives reported by the present sample of gymnasts are consistent with those outlined by Gould's (1987) model of youth sport withdrawal. This model theorizes that children's surface level reasons for participating in or withdrawing from sport reflect their underlying achievement goals. Empirical evidence supports this notion indicating that young athletes initially become involved and remain involved in sport largely due to intrinsic motives such as having fun, improving one's skill and fitness, and challenging one's ability (Ewing, Seefeldt, & Brown, 1996; Seefeldt et al., 1993). It has also been demonstrated that children may withdraw from sport because their intrinsic needs remain unfulfilled. This motive is typically reflected in statements such as 'not having fun,' 'being too bored,' and 'not getting enough playing time' (Burton & Martens, 1986; Petlichkoff, 1993).

Given that the assessment of participation motives within youth sport research typically occurs subsequent to dropout, it is not clear whether these motives contribute to sport dropout or, alternately, represent a retrospective artifact of withdrawal. The present study had the advantage of assessing the participation motives of athletes prior to their eventual withdrawal. This representative sample of competitive Australian gymnasts substantiates the notion that young athletes tend to engage in or withdraw from competitive sport largely as a function of their intrinsic and extrinsic participation motives. In particular, the motives of continuing gymnasts appeared to be more intrinsically based, whereas dropouts reported more extrinsic motives for their participation.

These findings may be linked to the capacity of young athletes to develop personally within the competitive sport setting. Duda and Nicholls (1992) state that intrinsically-involved athletes characteristically define their personal competence within the sport setting in terms of self-referenced standards such as effort, skill improvement, and learning. Extrinsically motivated athletes are more likely to assess their competence on the basis of norm-referenced criteria such as out-performing others and demonstrating superior ability with nominal effort. Because intrinsic involvement emphasizes self-referenced standards of personal ability, athletes adopting this perspective have a greater capacity to develop and maintain competence perceptions within the sport context. Based on these arguments, we would expect a relationship to exist between athletes' participation motives, their capacity for personal development within the sport context, and their tendency to dropout. Further work needs to assess the relative costs and benefits of competitive sport participation with regard to the overall personal development of young athletes. Research indicates that competitive youth sport has the potential of fostering both positive and negative personal outcomes among its participants. Although youth sport involvement has been reported to facilitate self-concept (Marsh, 1998), self-esteem (Kavanassu & Harnisch, 2000), body image (Miller & Levy, 1996), achievement-oriented attitudes (Curry, Snyder, Cook, Ruby, & Rehm, 1997), and general mental health (Steiner, McQuivey, Pavelski, Pitts, & Kraemer, 2000), other studies have found that youth athletes are lower than their non-participating peers in altruism (Blair, 1985), moral development (Shields & Bredemeier, 1995), and sportspersonship (Allison, 1982).

The governing bodies of Australian youth sport have advocated the general importance of youth sport participation *per se* and have considered sport attrition detrimental to the personal growth among young people. Recent attrition rates have also prompted criticism regarding the structure of competitive youth sport as well as concern over the seeming inadequacies of children who drop out. It is quite clear, however, that competitive youth sport attrition occurs for a variety of reasons. For example, it is quite possible that some of the dropouts in this study perceived a greater likelihood of enhancing their personal development in a setting other than competitive youth gymnastics. This attrition motive may, in fact, represent progress with respect to the child's developmental process rather than a limitation of the particular sport structure (Petlichkoff, 1996). As noted by Weiss & Ebbeck (1996), further research needs to clarify the particular role of sport involvement in the overall development of young athletes rather than assume that sport participation is necessarily in a child's best interest. With regard to the present study, such work could investigate the relative contribution of competitive gymnastics participation in facilitating young people's achievement of personal development goals such as autonomy, self-identity, creativity, and initiative.

Clearly, significant others such as coaches, parents, teachers, and teammates play a significant role in the formation of self-perceptions and participation attitudes among young athletes (Brustad, 1996; Kimiecik, Horn, & Shurin, 1996; Smoll, Smith, Barnett, & Everett, 1993; Weiss & Smith, 1999). It would seem important to develop a predictive model that would determine the contextual and dispositional factors that influence children's intrinsic and extrinsic participation motives as well as the conditions under which such dropout most hinders their progress towards personal development.

REFERENCES

Allison, M. T. (1982). Sportsmanship: Variations based on sex and degree of competitive experience. In A.O. Dunleavy, A.W. Miracle, & C.R. Rees (Eds.), *Studies in the sociology of sport* (pp. 153-165). Ft. Worth: Texas Christian University Press.

Athletic Footwear Association (1990). *American youth sports participation.* North Palm Beach, FL: American Footwear Association.

Australian Bureau of Statistics. (2001). *Children's participation in cultural and leisure activities, Australia* (ABS Publication No. 4901.0). Canberra, Australian Capital Territory: Author.

Australian Sports Commission. (1998). *AUSSSIE SPORT: Developing active young Australians* (ASC Publication No. 11-217). Belconnen, Australian Capital Territory: Author.

Australian Sports Commission. (1999). *The Australian Sports Commission: Beyond 2000* (ASC Publication No. 10-039). Belconnen, Australian Capital Territory: Author.

Blair, S. (1985). Professionalization of attitude toward play in children and adults. *Research Quarterly in Exercise and Sport, 56,* 82-83.

Brustad, R. J. (1996). Parental and peer influence on children's psychological development through sport. In F.L. Smoll & R.E. Smith (Eds.), *Children and youth in sport: A biopsychosocial perspective* (pp. 112-124). Madison, WI: Brown & Benchmark.

Burton, D. (1988). The dropout dilemma in youth sport: Documenting the problem and identifying solutions. In R. M. Malina (Ed.), *Young athletes: Biological, psychological, and educational perspectives* (pp. 245-266). Champaign, IL: Human Kinetics.

Burton, D. & Martens. R. (1986). Pinned by their goals: An exploratory investigation into why kids drop out of wrestling. *Journal of Sport Psychology, 8,* 183-197.

Clough, J., & Traill, R. (1992). *A mapping of participation rates in junior sport in the Australian Capital Territory.* Canberra, Australia: ACT Government Press.

Curry, L. A., Snyder, C. R., Cook, D. L., Ruby, B. L., & Rehm, M. (1997). Role of hope in academic and sport achievement. *Journal of Personality & Social Psychology, 73,* 1257-1267.

Duda, J. L. (1989). The relationship between goal perspectives and persistence and intensity among recreational sport participants. *Leisure Sciences,10,* 95-106.

Duda, J. L., & Nicholls, J. G. (1992). Dimensions of achievement motivation in schoolwork and sport. *Journal of Educational Psychology, 84,* 290-299.

Ewing, M. E., Seefeldt, V. D., & Brown, T. P. (1996*). Role of organized sport in the education and health of American children and youth: Background report on the role of sports in youth development.* New York: Carnegie Corporation of New York.

Fry, D. A., McClements, J. D., Sefton, J. M. (1981). *A report on participation in the Saskatoon Hockey Association.* Saskatoon, Canada: SASK Sport Press.

Gould, D. (1987). Understanding attrition in children's sport. In D. Gould, & M.R. Weiss (Eds.), *Advances in pediatric sport sciences* (Vol. 2, pp. 61-85). Champaign, IL: Human Kinetics.

Gould, D., Feltz, D., Horn, T., & Weiss, M. R. (1982). Reasons for discontinuing involvement in competitive youth swimming. *Journal of Sport Behavior, 5,* 155-165.

Gould, D., Feltz, D., & Weiss, M. R. (1985). Motives for participating in competitive youth swimming. *International Journal of Sport Psychology, 16,* 126-140.

Gould, D. & Medbery, R. (2000). Psychological issues. In S.J. Anderson & J.A. Sullivan (Eds.), *Care of the young athlete* (pp. 105-133). Elk Grove Village, IL: American Academy of Pediatrics Press.

Hall, H. (1990). *A social-cognitive approach to goal setting: The mediating effects of achievement goals and perceived ability.* Unpublished doctoral dissertation, University of Illinois.

Huberty, C. J., & Morris, J. D. (1989). Multivariate analysis versus multiple univariate analyses. *Psychological Bulletin, 105,* 302-308.

Kavussanu, M., & Harnisch, D. L. (2000). Self-esteem in children: Do goal orientations matter? *British Journal of Educational Psychology, 70,* 229-242.

Kimiecik, J. C., Horn, T. S., & Shurin, C. S. (1996). Relationships among children's beliefs, perceptions of their parents' beliefs, and their moderate-to-vigorous physical activity. *Research Quarterly for Exercise and Sport, 67,* 324-336.

Klint, K. A., & Weiss, M. R. (1986). Dropping in and dropping out: Participation motives of current and former youth gymnasts. *Canadian Journal of Applied Sport Sciences, 11,* 106-114.

Klint, K. A., & Weiss, M. R. (1987). Perceived competence and motives for participating in youth sports: A test of Harter's competence motivation theory. *Journal of Sport Psychology, 9,* 55-65.

Marsh, H. W. (1994). Sport motivation orientations: Beware of jingle-jangle fallacies. *Journal of Sport and Exercise Psychology, 16,* 365-380.

Marsh, H. W. (1998). Age and gender effects in physical self-concepts for adolescent elite athletes and nonathletes: A multicohort-multioccasion design. *Journal of Sport & Exercise Psychology, 20,* 237-259.

Maxwell, S. E. (1992). Recent developments in MANOVA applications. In B. Thompson (Ed.), *Advances in social science methodology* (pp. 137-168). Greenwich, CT: JAI Press.

Miller, J. L., & Levy, G. D. (1996). Gender role conflict, gender-typed characteristics, self-concepts, and sport socialization in female athletes and nonathletes. *Sex Roles, 35,* 111-122.

Petlichkoff, L. M. (1992). Youth sport participation and withdrawal: Is it simply a matter of fun? *Pediatric Exercise Science, 4,* 105-110.

Petlichkoff, L. M. (1993). Coaching children: Understanding the motivational process. *Sport Science Review, 2,* 48-61.

Petlichkoff, L. M. (1996). The drop-out dilemma in youth sports. In O. Bar-Or (Ed.), *The child and adolescent athlete: Encyclopaedia of sports medicine* (Vol. 6; pp. 418-430). Oxford, England: Blackwell Scientific Publications.

Robertson, I. (1981). *Children's perceived satisfactions and stresses in sport.* Paper presented at the Australian Conference on Health, Physical Education and Recreation Biennial Conference, Melbourne.

Robertson, I. (1992). *Children, AUSSIESPORT, and organised junior sport.* Canberra, Australia: Australian Sports Commission.

Seefeldt, V., Ewing, M., & Walk, S. (1993). *Overview of youth sports programs in the United States.* Unpublished manuscript commissioned by the Carnegie Council On Adolescent Development.

Seifriz, J., Duda, J. L., & Chi, L. (1992). The relationship between perceived motivational climate to achievement-related affect and cognitions in basketball. *Journal of Sport & Exercise Psychology, 14,* 375-391.

Shields, D. L., & Bredemeier, B. J. (1995). *Character development and physical activity.* Champaign, IL: Human Kinetics.

Smoll, F. L., Smith, R. E., Barnett, N. P., & Everett, J. J. (1993). Enhancement of children's self-esteem through social support training for youth sport coaches. *Journal of Applied Psychology, 78,* 602-610.

Steiner, H., McQuivey, R. W., Pavelski, R., Pitts, T., & Kraemer, H. (2000). Adolescents and sports: Risk or benefit? *Clinical Pediatrics, 39,* 161-166.

Tabachnick, B. G., & Fidell, L. S. (1996). *Using multivariate statistics* (3rd ed.). Harper Collins.

Weiss, M. R., & Ebbeck, (1996). Self-esteem and perceptions of competence in youth sports: Theory, research, and enhancement strategies. In O. Bar-Or (Ed.), *The child and adolescent athlete: Encyclopedia of sports medicine: Vol. 6* (pp. 364-382). Oxford, England: Blackwell Scientific.

Weiss, M. R., & Glenn, S. D. (1992). Psychological development and females' sport participation: An interactional perspective. *Quest, 44,* 138-157.

Weiss, M. R., & Petlichkoff, L. M. (1989). Children's motivation for participation in and withdrawal from sport: Identifying the missing links. P*ediatric Exercise Science, 1,* 195-211.

Weiss, M. R., & Smith, A. L. (1999). Quality of youth sport friendships: Measurement development and validation. *Journal of Sport & Exercise Psychology, 24,* 145-166.

Weitzer, J. E. (1989). *Childhood socialization into physical activity: Parental roles in perceptions of competence and goal orientation.* Unpublished master's thesis, University of Wisconsin, Milwaukee.

The Internal Consistency Reliability and Construct Validity of the German Translation of the Oxford Happiness Inventory

Christopher Alan Lewis
University of Ulster at Magee College, Northern Ireland

Leslie J. Francis
University of Wales, United Kingdom

Hans-Georg Ziebertz
University of Würzburg, Germany

German translations of the Oxford Happiness Inventory and the short form Revised Eysenck Personality Questionnaire were completed by 331 students in Germany. The findings confirm the internal reliability of the German translation of the Oxford Happiness Inventory and support the construct validity of the instrument according to which 'happiness is a thing called stable extraversion.'

The Oxford Happiness Inventory (OHI) was developed by Argyle, Martin and Crossland (1989) to provide a general measure of happiness. Drawing on earlier discussion by Argyle and Crossland (1987), they suppose that happiness comprises three main components: the frequency and degree of positive affect or joy; the average level of satisfaction over a period; and the absence of negative feelings, such as depression and anxiety. Working from this definition, they developed the Oxford Happiness Inventory by reversing the 21 items of the Beck Depression Inventory (Beck, Ward, Mendelson, Hock, & Erbaugh, 1961) and adding 11 further items to cover aspects of subjective wellbeing not so far included. Three items were subsequently dropped, leading to a 29-item scale. Each item invited the respondents to select one of four options. The incremental steps in these items are defined as follows: unhappy or mildly depressed (e.g., "I do not feel happy"); a low level of happiness (e.g., "I feel fairly happy"); a high level of happiness (e.g., "I am very happy"); and manic (e.g., "I am incredibly happy").

Argyle, Martin and Crossland (1989) reported an internal reliability of .90 using alpha (Cronbach, 1951) and a seven week test-retest reliability of

---

Author info: Correspondence should be sent to: Leslie J. Francis, Welsh National Centre for Religious Education, University of Wales, Bangor Gwynedd LL57 2PX UK

North American Journal of Psychology, 2002, Vol. 4, No. 2, 211-220.

© NAJP

.78. Argyle and Lu (1990b) reported a test-retest reliability of .81 after four months, and Valiant (1993) reported a test-retest reliability of .83 after approximately six weeks. Other studies have reported similarly high alpha coefficients; for example, .87 by Furnham and Brewin (1990), .90 by Brebner, Donaldson, Kirby and Ward (1995), .89 by Robbins and Francis (1996), .90 by Hills and Argyle (1998b), .92 by Brebner (1998), .89, .89, .90 and .90 in the four samples studied by Francis, Brown, Lester and Philipchalk (1998), .92 by Francis (1999), .92 by French and Joseph (1999), .90, .91 and .92 in three samples reported by Francis, Jones and Wilcox (2000), .89 by Francis and Robbins (2000), and .86 by Chan and Joseph (2000). Using an 11-item short-form of the *Oxford Happiness Inventory*, Noor (1997) reported an alpha value of .83.

Evidence of the construct validity of the *Oxford Happiness Inventory* has been provided by a set of nine studies that have located scores on this inventory within the context of Eysenck's model of personality. According to the model, individual differences can be most economically and adequately summarised in terms of the three orthogonal dimensions of extraversion, neuroticism and psychoticism (Eysenck & Eysenck, 1985). These studies, conducted among samples of 131 undergraduates in Oxford (Argyle & Lu, 1990a), 101 students in London (Furnham & Brewin, 1990), 114 adults in Oxford (Lu & Argyle, 1991), 581 Chinese adults living in Taiwan (Lu, 1995), 95 students in Australia (Brebner, Donaldson, Kirby, & Ward, 1995), a total of 1,076 students in the UK, USA, Australia and Canada (Francis, Brown, Lester, & Philipchalk, 1998), 456 students in Wales (Francis, 1999), 120 students in the United Kingdom (Furnham & Cheng, 1999), 233 English speaking young people mainly recruited from the final year at school (Furnham & Cheng, 2000) and 107 students in England (Chan & Joseph, 2000) have demonstrated that happiness is correlated positively with extraversion and correlated negatively with neuroticism. A positive relationship between happiness and extraversion and a negative relationship between happiness and neuroticism using the NEO Five Factor Inventory (Costa & McCrae, 1989) has been demonstrated by Furnham and Cheng (1997) among 83 young adults and by Brebner (1998) among 86 Adelaide students. A positive relationship between happiness and extraversion using the *Myers-Briggs Type Indicator* (Myers & McCaulley, 1985) has been demonstrated by Francis and Jones (2000) among 284 adult churchgoers. These findings are consistent with Eysenck's (1983) definition of the construct happiness:

> Happiness is a thing called stable extraversion ... [and since] the positive affect in happiness seemed to be related to easy sociability, with a natural, pleasant interaction with other people, ...then it only makes sense that happiness can be associated with extraversion. Similarly, if worries and

> anxieties make up negative affect [rather than] happiness, it can easily be seen that instability and neuroticism are also connected to unhappiness.

A series of studies employing the *Oxford Happiness Inventory* has begun to map the correlates of this operational definition of happiness (Argyle, Martin, & Lu, 1995). For example, positive predictors of happiness have been identified as social competence (Argyle & Lu, 1990b), self-esteem, social skills and cooperativeness (Lu & Argyle, 1991), satisfaction with relationships with people from whom support had been received (Lu & Argyle, 1992), engagement in a serious leisure activity (Lu & Argyle, 1994). Other studies have identified positive predictors of happiness as internal locus of control (Noor, 1993), social support (Lu, 1995), religiosity (Argyle & Hills, 2000; Francis, Jones, & Wilcox, 2000; Francis & Lester, 1997; Francis & Robbins, 2000; Robbins & Francis, 1996), intensity of musical experience (Hills & Argyle, 1998a), participation in sports (Hills & Argyle, 1998b), low levels of depression (Joseph & Lewis, 1998; Joseph, Harwood, Lewis & McCollam, 2002), self-actualisation and purpose in life (French & Joseph, 1999), and life regard, self esteem, life satisfaction and affiliative tendency (Hills & Argyle, 2001). Lu and Argyle (1993) found an inverse relationship between happiness and the total time spent watching television. Rim (1993) found a relationship between happiness and coping styles. Both men and women who recorded high happiness scores were high on mapping coping style and low on substitution coping style.

Lu and Shih (1997) adapted the Oxford Happiness Inventory for use among Chinese people. One of the original 29 items encountered problems in translation. A further 20 items were added to cover distinctive aspects of Chinese happiness. The new 48-item *Chinese Happiness Inventory* (CHI) achieved an alpha value of 0.94. Alongside the *Eysenck Personality Questionnaire* (Eysenck & Eysenck, 1975) happiness was positively correlated with extraversion and negatively correlated with neuroticism as in other Western studies. These findings were replicated in further studies using the Chinese Happiness Inventory as reported by Lu, Shih, Lin and Ju (1997) and Lu and Lin (1998).

Lu, Gilmour, Kao, Weng, Hu, Chern, Huang, and Shih (2001) produced a brief 20-item form of the Chinese Happiness Inventory for use among a Taiwanese sample. They reported an alpha coefficient of .90. This study did not include the Eysenck personality dimensions.

Furnham and Cheng (1999) translated the *Oxford Happiness Inventory* into Japanese. In a study conducted among 70 female and 58 male Japanese speaking students attending a professional school, they found that the 29-item scale produced an alpha coefficient of .86. Alongside the 90-item

Eysenck Personality Questionnaire happiness was positively correlated with extraversion but independent of both neuroticism and psychoticism.

Francis and Katz (2000) translated the *Oxford Happiness Inventory* into Hebrew. In a study conducted among 298 female Hebrew speaking undergraduates attending Bar-Ilan University they found that all 29 items functioned satisfactorily, producing an alpha coefficient of .85. Once again, alongside the short form *Revised Eysenck Personality Questionnaire* happiness was positively correlated with extraversion and negatively correlated with neuroticism.

Neto (2001) translated the *Oxford Happiness Inventory* into Portuguese. In a study conducted among 130 female and 41 male Portuguese speaking undergraduate students, he found that each of the 29 scale items correlated at least .30 with the sum of all the other items, producing an alpha coefficient of .90. This study did not include the Eysenck personality dimensions.

The aim of the present study was to extend the current cross cultural work on the *Oxford Happiness Inventory*. We developed a German translation of this instrument and subsequently examined its internal consistency reliability and construct validity in terms of Eysenck's dimensional model of personality among a sample of German-speaking students.

METHOD

Measures

The *Oxford Happiness Inventory* (Argyle, Martin, & Crossland, 1989) is a 29-item instrument. Each item invites the respondent to chose one of four sentences constructed to reflect incremental steps defined as: unhappy or mildly depressed, a low level of happiness, a high level of happiness, and mania. Each item was translated into German and then back-translated into English in order to check the accuracy and unambiguity of the translation.

The short form *Revised Eysenck Personality Questionnaire* (Eysenck, Eysenck, & Barrett, 1985) is a 48-item instrument composed of four 12-item measures of extraversion, neuroticism, psychoticism and a lie scale. Each item is assessed on a two-point scale: yes and no. Data on the reliability of the German translation of this instrument have been provided by Francis, Ziebertz and Lewis (2002), who reported satisfactory alpha coefficients.

Participants and Procedure

Both instruments were completed during class time by a sample of 311 German speaking undergraduate students attending the University of Würzburg. Of the total respondents, 19 were under the age of twenty, 298 were in their twenties, 10 in their thirties and four were aged forty or over; 204 were female and 127 were male. The data were analysed using the

Statistical Package for the Social Sciences (SPSS Inc., 1988) employing the reliability, descriptives and correlation procedures.

RESULTS

Table one presents the item rest of test correlation coefficients for each of the 29 items of the German translation of the *Oxford Happiness Inventory*. These figures confirm that each item makes a positive contribution to the overall scale score. The scale achieved an alpha coefficient of .86.

TABLE 1 The Oxford Happiness Inventory: Item to Rest of Test Correlations

Inventory items	r
I am incredibly happy	.51
I feel sure that the future is overflowing with hope and promise	.41
I am totally satisfied with everything	.37
I feel that I am in total control of all aspects of my life	.38
I feel that life is overflowing with rewards	.53
I am delighted with the way I am	.50
I always have a good influence on events	.38
I love life	.58
I am intensely interested in other people	.20
I can make all decisions very easily	.27
I feel able to take anything on	.37
Nowadays I always wake up feeling more rested than I used to	.29
I feel I have boundless energy	.46
The whole world looks beautiful to me	.41
I have never felt so mentally alert as I do nowadays	.39
I feel on top of the world	.39
I love everybody	.31
All past events seem extremely happy	.37
I am constantly in a state of joy and elation	.53
I have done everything I ever wanted	.38
My time is perfectly organised so that I can fit in all things I want to do	.33
I always have fun with other people	.46
I always have a cheerful effect on others	.38
My life is totally meaningful and purposive	.48
I am always committed and involved	.26
I think the world is an excellent place	.31
I am always laughing	.44
I think I look exceptionally attractive	.41
I am amused by everything	.48

The four scales of the short form Revised Eysenck Personality Questionnaire achieved the following alpha coefficients: extraversion, .85; neuroticism, .81; psychoticism, .42; lie scale, .64. The lower alpha coefficient for the psychoticism scale is consistent with the known difficulties in measuring this dimension of personality (Francis, Philipchalk, & Brown, 1991).

TABLE 2 The Oxford Happiness Inventory: Mean Scale Scores by Sex

Scale	Male		Female		
	Mean	*sd*	Mean	*sd*	*T*
Oxford Happiness Inventory	41.55	10.8	43.05	10.2	1.20
EPQR-S extraversion	7.39	3.3	7.94	3.2	1.42
EPQR-S neuroticism	5.10	3.1	6.14	3.2	2.89*
EPQR-S psychoticism	2.67	1.8	2.59	1.5	0.39
EPQR-S lie scale	3.81	2.2	4.02	2.3	0.83

* $p < .01$

Table two presents the mean scale scores of happiness, extraversion, neuroticism, psychoticism and the lie scale for males and for females separately. These data show that the only significant difference between the sexes occurs in relationship to neuroticism scores. The finding that females record higher neuroticism scores than males is consistent with the consensus of the international literature as summarised by Francis (1993).

Table three presents the Pearson correlation coefficients between the German translation of the *Oxford Happiness Inventory*, extraversion, neuroticism, psychoticism and lie scale scores. These statistics confirm that happiness is related positively with extraversion and negatively with neuroticism, but is unrelated to psychoticism.

TABLE 3 Correlation Matrix

	Lie Scale	Psychoticism	Neuroticism	Extraversion
Happiness	.19**	-.05	-.46**	.46**
Extraversion	-.01	.09	-.27**	
Neuroticism	-.15*	.07		
Psychoticism	-.15*			

** $p < .001$ * $p < .01$

DISCUSSION

Three main conclusions emerge from the findings of this study. First, the data support the internal consistency reliability of the German translation of the Oxford Happiness Inventory among undergraduate students in Germany. Although four of the item rest of test correlations fall below the threshold of 0.3, the alpha coefficient is a more than adequate indicator of the overall item homogeneity of the scale. Given the balance of items within the scale to reflect Argyle and Crossland's (1987) construct of happiness, it is not recommended that the weaker items should be dropped to form a shorter scale.

Second, in line with a number of other studies, no significant sex differences were found in mean scores on the *Oxford Happiness Inventory* (Argyle & Lu, 1990a; Francis, Brown, Lester, & Philipchalk, 1998; Furnham & Brewin, 1990; Lu & Argyle, 1991, 1992, 1993).

Third, the data support the construct validity of the German translation of the *Oxford Happiness Inventory* among undergraduate students in Germany. Within the context of Eysenck's dimensional model of personality, happiness in Germany is clearly 'a thing called stable extraversion', as was also found to be the case in the UK, USA, Australia and Canada by Francis, Brown, Lester and Philipchalk (1998); in Australia by Brebner, Donaldson, Kirby and Ward (1995); in China by Lu and Shih (1997), Lu, Shih, Lin and Ju (1997) and Lu and Lin (1998); and in Israel by Francis and Katz (2000).

Further replications, using the *Oxford Happiness Inventory* alongside Eysenck's dimensional model of personality, are now required to check these findings among other samples in Germany, including both school children and adults. Further support for the validity of the *Oxford Happiness Inventory* could be provided by including additional indices of the three component parts of happiness identified by Argyle and Crossland (1987), namely the frequency and degree of positive affect or joy, the average level of satisfaction over a period, and the absence of negative feelings such as depression and anxiety.

REFERENCES

Argyle, M., & Crossland, J. (1987). Dimensions of positive emotions. *British Journal of Social Psychology, 26*, 127-137.

Argyle, M., & Hills, P. (2000). Religious experiences and their relations with happiness and personality. *International Journal for the Psychology of Religion, 10*, 157-172.

Argyle, M., & Lu, L. (1990a). The happiness of extraverts. *Personality and Individual Differences, 11*, 1011-1017.

Argyle, M., & Lu, L. (1990b). Happiness and social skills. *Personality and Individual Differences, 11*, 1255-1261.

Argyle, M., Martin, M., & Crossland, J. (1989). Happiness as a function of personality and social encounters. In J.P. Forgas & J.M. Innes (eds), *Recent advances in social psychology: An international perspective* (pp. 189-203). North Holland: Elsevier Science Publishers.

Argyle, M. Martin, M., & Lu, L. (1995). Testing for stress and happiness: the role of social and cognitive factors. In C.D. Spielberger & J. Brebner (eds), *Testing for* *stress and happiness: The role of social factors* (pp. 173-187). New York: Taylor and Francis.

Beck, T., Ward, C.H., Mendelson, M., Hock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry*, *7*, 158-216.

Brebner, J. (1998). Happiness and personality. *Personality and Individual Differences*, *25*, 279-296.

Brebner, J., Donaldson, J., Kirby, N., & Ward, L. (1995). Relationships between happiness and personality. *Personality and Individual Differences*, *19*, 251-258.

Chan, R., & Joseph, S. (2000). Dimensions of personality, domains of aspiration, and subjective well-being. *Personality and Individual Differences*, *28*, 347-354.

Costa, P.T. jr, & McCrae, R. (1989). *The NEO-PI/NEO-FFI manual supplement*. Odessa, Florida: Psychological Assessment Resources.

Cronbach, L.J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*, 297-334.

Eysenck, H.J. (1983). *I do: Your guide to a happy marriage*. London: Century.

Eysenck, H.J., & Eysenck, M.W. (1985). *Personality and individual differences: A natural science approach*. New York: Plenum Press.

Eysenck, H.J., & Eysenck, S.B.G. (1975). *Manual of the Eysenck Personality Questionnaire (adult and junior)*. London: Hodder and Stoughton.

Eysenck, S.B.G., Eysenck, H.J., & Barrett, P. (1985). A revised version of the psychoticism scale. *Personality and Individual Differences*, *6*, 21-29.

Francis, L.J. (1991). The dual nature of the EPQ lie scale among college students in England. *Personality and Individual Differences*, *12*, 1255-1260.

Francis, L.J. (1993). The dual nature of the Eysenckian neuroticism scales: a question of sex differences? *Personality and Individual Differences*, *15*, 43-59.

Francis, L.J. (1999). Happiness is a thing called stable extraversion: a further examination of the relationship between the Oxford Happiness Inventory and Eysenck's dimensional model of personality and gender. *Personality and Individual Differences*, *26*, 5-11.

Francis, L.J., Brown, L.B., Lester, D., & Philipchalk, R. (1998). Happiness as stable extraversion: a cross-cultural examination of the reliability and validity of the Oxford Happiness Inventory among students in the UK, USA, Australia and Canada. *Personality and Individual Differences*, *24*, 167-171.

Francis, L.J., & Jones, S.H. (2000). Psychological type and happiness: a study among adult churchgoers. *Journal of Psychological Type*, *54*, 36-41.

Francis, L.J., Jones, S.H., & Wilcox, C. (2000). Religiosity and happiness: during adolescence, young adulthood and later life. *Journal of Psychology and Christianity*, *19*, 245-257.

Francis, L.J., & Katz, Y. (2000). The internal consistency reliability and validity of the Hebrew translation of the Oxford Happiness Inventory. *Psychological Reports*, *87*, 193-196.

Francis, L.J., & Lester, D. (1997). Religion, personality and happiness. *Journal of Contemporary Religion*, *12*, 81-86.

Francis, L.J., Philipchalk, R., & Brown, L.B. (1991). The comparability of the short form EPQ-R with the EPQ among students in England, the USA, Canada and Australia. *Personality and Individual Differences*, *12*, 1129-1132.

Francis, L.J., & Robbins, M. (2000). Religion and happiness: a study in empirical theology. *Transpersonal Psychology Review, 4* (2), 17-22.

Francis, L.J., Ziebertz, H.-G., & Lewis, C.A. (2002). The Short-Form Revised Eysenck Personality Questionnaire (EPQR-S): a German edition, (in press).

French, S., & Joseph, S. (1999). Religiosity and its association with happiness, purpose in life, and self-actualisation. *Mental Health, Religion and Culture, 2,* 117-120.

Furnham, A., & Brewin, C.R. (1990). Personality and happiness. *Personality and Individual Differences, 11,* 1093-1096.

Furnham, A., & Cheng, H. (1997). Personality and happiness. *Psychological Reports, 80,* 761-762.

Furnham, A., & Cheng, H. (1999). Personality as predictors of mental health and happiness in the East and West. *Personality and Individual Differences, 27,* 395-403.

Furnham, A., & Cheng, H. (2000). Lay theories of happiness. *Journal of Happiness Studies, 1,* 227-246.

Hills, P., & Argyle, M. (1998a). Musical and religious experiences and their relationship to happiness. *Personality and Individual Differences, 25,* 91-102.

Hills, P., & Argyle, M. (1998b). Positive moods derived from leisure and their relationship to happiness and personality. *Personality and Individual Differences, 25,* 523-535.

Hills, P., & Argyle, M. (2001). Happiness, introversion-extraversion and happy introverts. *Personality and Individual Differences, 30,* 295-608.

Joseph, S., Harwood, J., Lewis, C.A., & McCollam, P. (2002). Assessing subjective well-being: a short form of the Depression-Happiness Scale (SDHS). *Journal of Counseling Psychology,* (in press).

Joseph, S., & Lewis, C.A. (1998). The Depression-Happiness Scale: reliability and validity of a bipolar self-report scale. *Journal of Clinical Psychology, 54,* 537-544.

Lu, L. (1995). The relationship between subjective well-being and psychosocial variables in Taiwan. *Journal of Social Psychology, 135,* 351-357.

Lu, L., & Argyle, M. (1991). Happiness and cooperation. *Personality and Individual Differences, 12,* 1019-1030.

Lu, L., & Argyle, M. (1992). Receiving and giving support: effects on relationships and well-being. *Counseling Psychology Quarterly, 5,* 123-133.

Lu, L., & Argyle, M. (1993). TV watching, soap opera and happiness. *Kaohsiung Journal of Medical Sciences, 9,* 501-507.

Lu, L., & Argyle, M. (1994). Leisure satisfaction and happiness: a function of leisure activity. *Kaohsiung Journal of Medical Sciences, 10,* 89-96.

Lu, L., Gilmour, R., Kao, S.-F., Weng, T.-H, Hu, C.-H., Chern, J.-G., Huang, S.-W., & Shih, J.-B. (2001). Two ways to achieve happiness: when the East meets the West. *Personality and Individual Differences, 30,* 1161-1174.

Lu, L., & Lin, Y.Y. (1998). Family roles and happiness in adulthood. *Personality and Individual Differences, 25,* 195-207.

Lu, L., & Shih, J.B. (1997). Personality and happiness: is mental health a mediator? *Personality and Individual Differences, 22,* 249-256.

Lu, L., Shih, J.B., Lin, Y.Y., & Ju, L.S. (1997). Personal and environmental correlates of happiness. *Personality and Individual Differences*, *23*, 453–462.

Myers, I.B., & McCaulley, M.H. (1985). *Manual: A guide to the development and use of the Myers-Briggs Type Indicator*. Palo Alto, California: Consulting Psychologists Press.

Neto, F. (2001). Personality predictors of happiness. *Psychological Reports*, *88*, 817-824.

Noor, N.M. (1993). Work and family roles in relation to women's well-being, unpublished DPhil dissertation, University of Oxford.

Noor, N.M. (1997). Work and family roles in relationship to women's well-being: the role of negative affectivity. *Personality and Individual Differences*, *23*, 487-499.

Rim, Y. (1993). Happiness and coping styles. *Personality and Individual Differences*, *14*, 617-618.

Robbins, M., & Francis, L.J. (1996). Are religious people happier? A study among undergraduates. In L.J. Francis, W.K. Kay and W.S. Campbell (eds), *Research in Religious Education* (pp. 207-217). Leominster: Gracewing.

SPSS Inc. (1988). *SPSSX user's guide*. New York: McGraw-Hill

Valiant, G.L. (1993). Life events, happiness and depression: the half empty cup. *Personality and Individual Differences*, *15*, 447-453.

Body-image Assessment: A Review and Evaluation of a New Computer-aided Measurement Technique

Salvatore Cullari, Michelle Vosburgh, Amber Shotwell,
Julien Inzodda, & Wendi Davenport.
Lebanon Valley College

A new procedure for assessing body-image was investigated in this study. It appears that digitally distorted images of subjects are a viable measurement technique for body-image assessment. The current technique appears to solve some of the problems associated with the use of silhouettes and will probably prove to be much more effective than silhouettes with younger subjects. It also allows the magnitude of body-image distortion to be more easily and accurately measured. The current study replicated many of the common findings of previous studies including significant gender differences in body-image perceptions, dissatisfaction and distortion. The results of this study suggest that the etiology of body-image distortion in both males and females may be similar.

The topics of body-image assessment and body-image disturbance have been among the most prominent topics in the eating disorder literature for many years (Thompson,1996). One of the reasons for this is that both body-image dissatisfaction and distortion appear to be linked to eating disorders (Cullari, Rohrer, & Bahm, 1998). Furthermore, although results have not been uniform, both of these appear to influence self-concept and self-esteem, at least in females (Folk, Pedersen & Cullari, 1993). In recent years, the following findings have been consistently replicated.

Body-image distortion occurs in both men and women and in individuals with and without eating disorders (Cullari & Trubilla, 1989). In general, women tend to view themselves as heavier than they really are, while men tend to view themselves as the same or lighter than they really are (Fallon & Rozin, 1985). Body dissatisfaction in women starts at a very young age, which may be around the start of puberty or even younger (Rauste-von Wright, 1989; Folk, et al., 1993). About 75% of young women between the ages of 16 and 24 are dissatisfied with some aspect of their body. In contrast, only about 25% of young men of a similar age group are dissatisfied, although this latter figure may be growing (Cullari et al., 1998; Fallon & Rozin, 1985). In both men and

---

Author info: Correspondence should be sent to: Salvatore Cullari, Ph.D, Lebanon Valley College, Annville, PA 17003, or Cullari@lvc.edu.

North American Journal of Psychology, 2002, Vol. 4, No. 2, 221-232.
© NAJP

women, body-image dissatisfaction tends to precede body-image distortion and probably eating disorders as well (Cullari et al., 1998).

Although Americans spend about 50 billion dollars a year in an attempt to lose weight, 97% of all diets fail within a year (Garner, 1997). In most cases, people on diets gain the same weight back or more. Typically, people start diets because they are dissatisfied with their body weight. Studies have found that young womens' ideal weight is significantly less than their real weight. On the other hand, most young mens' ideal weight is about the same as their real weight (Cullari et al., 1998; Garner, 1997; Folk et al., 1993). As a result, women are much more likely to be on a weight reduction diet than men (Heatherton, Mahamedi, Striepe, Field, & Keel, 1997). For example, about 40% of girls under age 10 have already been on a weight reduction diet (Ogden & Mundray, 1996). Furthermore, an estimated 5 to 8 million American women have eating disorders and many eating disorders begin after a weight reduction diet (Garner, 1997).

Many researchers maintain that the ideal body image for both men and women is largely determined by social norms, which in turn are often transmitted by mass media (Heinberg & Thompson, 1995). For example, many young women use models in fashion magazines as the reference point for what they themselves should look like and believe that their health would be better if their bodies were similar to those of fashion models. However, over the last 30 years, female models in magazines, television and the movies have become significantly thinner (Kalodner, 1997; Rabak-Wagener, Eickhoff-Shemel, & Kelly-Vance, 1998), while the weight of many young women has actually increased.

The average adult female model in fashion magazines is about 15% less than their expected weight given their height and many female fashion models meet the DSM-IV-TR *weight* criteria for anorexia nervosa (Kalodner, 1997). At the same time, some studies suggest that exposure to thin female models leads to stress, depression, guilt and shame for young women who are dissatisfied with their own bodies (Pinhas, Toner, Ali, Garfinkel, & Stuckless, 1999). Similarly, it appears that exposure to thin female models leads to increased body-dissatisfaction and increased body distortion in women who already have eating disorders (Ogden & Mundray, 1996). To compound the problem, popular magazines for young women have 10 times the number of advertisements for weight reduction diets and exercise than popular magazines for young men (Wegner, Hartmann, & Geist, 2000).

Anderson (1988) estimates that the average American is bombarded with over 1,000 advertisements per day. Most of these ads both promote and link sexuality with beauty and extreme thinness in women. While many women models are quite thin, sexual attractiveness in males is often equated with average weight and increased muscularity. In this case, many women come to the mistaken belief that in order to be sexually attractive, they most be very thin. A likely result of this media influence is that most young women want to lose weight, while most young men want to gain weight or at least gain muscle mass. In order to reach their unrealistic ideal, women must undergo extreme dieting and exercise or engage in unhealthy or perhaps deadly eating disordered behaviors, while the ideal for men is much more reachable and perhaps even healthier. The irony is that while men tend to prefer women that are thin, their preference is not as thin as most young women estimate it to be (Fallon & Rozin, 1985).

Despite all of these findings, the field has not been free of controversy. One of the most important ones is the issue of how best to measure body image. Currently over 40 instruments for the measurement of body image exist (Thompson, Altabe, Johnson, & Stormer, 1994). These generally can be broken down into three categories: figure preferences (Stunkard, Sorenson, & Schlusinger, 1983); video projection techniques (Gleghorn, Penner, Powers, & Schulman, 1987); and questionnaires (Thompson, et al., 1994). Although many of these measurement techniques have been found to be valid and reliable (Gleghorn, et al., 1987; Thompson, et al., 1994), the sheer number of these different measurements makes it difficult to draw meaningful research generalizations, and may explain some of the inconsistencies observed in the literature.

One of the most often used methods of measuring body dissatisfaction and body distortion has been the use of silhouettes. While this method has been found to be generally satisfactory, some problems have been noted with the procedure itself and its use with young subjects (Byrne & Hills, 1996; Gardner, Friedman, & Jackson, 1998). For example, the silhouettes themselves are naked figure drawings, which do not look very realistic. Thus some subjects have problems trying to match the way they perceive themselves to the figures. In addition, the size differences between each adjacent figure drawing as a whole and the different body regions (e.g., chest and waist) are not uniform (see Gardner, et al., 1998). Therefore data interpretation is difficult. In a similar vein, quantifying the magnitude of body image distortion or body image dissatisfaction in a reliable manner is not possible with the use of figure drawings.

One of the problems stemming from the use of silhouettes with children and adolescents is that the original silhouettes drawings (e.g., Stunkard, et al., 1983) portrayed adults, and were not suitable for younger ages. Although silhouette drawings of children and adolescents have been developed (Byrne & Hills, 1996), the variability of the body size of children between age groups and even within the same age group is so great that it is almost impossible to develop standard sized models.

One of the purposes of the current study is to evaluate a new and relatively simple procedure to measure body image assessment that may eliminate some of the shortcomings of the silhouette method. In addition, this method was used in an attempt to replicate some of the previous research findings.

METHOD

Participants

Eighty (twenty-two male and fifty-eight female) college students at a small liberal arts college in the Northeast volunteered for the study. The participants ranged in age from 18 to 42, with a mean of 19.7 and SD of 3.4. Ninety-five percent of the subjects were white. Five subjects (one male and four female) had reported a previous (but not current) history of eating disorders.

Material and Procedure

A full figure (head to feet) digital color photograph of each subject was taken wearing a tee shirt and shorts. This image was then distorted using Adobe Photoshop® so that the subject's real image was increased up to 50% larger and reduced to 50% smaller than actual size. This was done at 10% intervals so that there were 5 images of each subject that were smaller than their actual size and 5 images that were larger than actual size. These ten distorted images plus the subject's original undistorted digital image were shown on a random order over a 15-inch color computer monitor. There were no time limits for viewing these eleven pictures, and participants could view each photo as often as they wanted. The participants were then asked to identify one photograph that was: 1) their real image; 2) their ideal image; 3) the image the subject believed was most attractive; 4) the image the subject believed would be rated most attractive by the opposite sex; 5) the image the subject believed would be rated most attractive by the same sex. These five choices could be the same or different photographs. On average, it took each subject between 15 and 20 minutes to complete these selections.

Similar to a number of previous studies using silhouettes, the difference between the subject's rating (on a 11 point scale) of their undistorted photograph and the one the subject perceived as their own image was used as a measure of body-image distortion. The difference between the subject's undistorted photograph and the one they chose as their ideal was used as a measure of body-image dissatisfaction.

In addition to basic demographic data, each participant was asked to report both real and ideal body weight. The difference between these two was used as a measure of body *weight* dissatisfaction.

The following measures were also used:

Fear of Negative Evaluation (FNE, Watson & Friend, 1969). This is a 30-item instrument designed to measure the fear of receiving negative evaluations from others (example: I am afraid that others will not approve of me).

Body Image Avoidance Questionnaire (BIAQ, Rosen, Srebnik, Saltzberg, & Wendt, 1991). This is a 19-item instrument designed to measure various avoidance tendencies related to body image disturbance (examples: I wear baggy clothes; I don't wear revealing clothes).

Multidimensional Body-self Relations Questionnaire (MBRQ, Cash & Pruzinsky, 1990). This is a 69-item instrument designed to measure subjective aspects of body image. The following subscales were used: appearance evaluation (AE, 7 items; example: my body is sexually appealing); appearance orientation (AO, 12 items; example: before going out, I usually spend a lot of time getting ready); fitness evaluation (FE, 3 items; example: I do poorly in physical sports or games); fitness orientation (FO, 13 items; example: I would pass most physical fitness tests); body areas satisfaction (BAS, 9 items; example: satisfaction of height, weight and overall appearance); subjective weight (SW, 2 items; example: from looking at me, most people would think I'm underweight, normal or overweight); weight preoccupation (WP, 4 items; example, I constantly worry about being or becoming fat).

Eating Disorder Inventory-2, (Garner, 1991). Only the 7-item Drive for Thinness (example, I think about dieting) and 9-item Body Dissatisfaction (example, I think my hips are too big) subscales were used.

RESULTS

Note that due to the large number of correlations and t-tests included in this study, we used the .01 level of significance in all statistical analyses. Because of this, some of the findings that were significant at the .05 level of significance may warrant further study.

Body Weight

Twenty participants (7 males and 13 females) were on weight reduction diets. This difference was not statistically significant. The mean weight and height for men was 179.3 pounds (*SD*=36.0) and 70.3 inches (*SD*=3.5) respectively. As a group, the male subjects were about 15 pounds heavier than the body height-weight range (151-163 pounds) suggested by the Metropolitan Life Insurance height and weight tables (Metropolitan Life Insurance, 1999). For males, there was no significant difference between their real weight (M =179.3) and ideal weight (M = 177.0; SD=30.2; repeated measure t =.70, $p > .05$), or the weight they perceived was the most attractive to the opposite sex (M = 168.2; SD=27.7; t = 1.6, $p > .05$). The mean weight and height for females was 136.8 pounds (SD=23.6) and 65.1 inches (SD=2.9) respectively. As a group, the mean weight of the females (136.8 pounds) was well in the range (127-141) suggested by the Metropolitan Life Insurance height and weight tables. However, for females, there was a significant difference between their real weight (M=136.8) and ideal weight (M=124.9; SD=17.1; repeated measure t=7.5, $p < .01$), as well as their real weight and the weight they perceived was the most attractive to the opposite sex (M=123.0, SD= 15.8; t = 7.9, $p < .01$).

Body-image perception, distortion and satisfaction: Picture ratings

There were significant gender differences in body-image perceptions of photographs (mean gender difference =.8; independent t = -2.7, $p < .01$). As a group, men perceived themselves as lighter than they really were (mean rating difference = .2), but this difference was not statistically significant (repeated measure t = .75, $p > .05$). Women perceived themselves as heavier than their real image, and this difference was statistically significant (mean rating difference = -.6, repeated measure t = -4.2, $p < .01$). In a similar vein, there were significant gender differences between perceived real and ideal figures (body-image dissatisfaction). Men choose pictures of their ideal selves that were heavier than their perceived real (M difference = -.50, repeated measure t = -.78, $p > .05$), but this difference was not statistically significant. Females chose ideal figures that were statistically smaller than their perceived real self (M difference = +1.85, repeated measure t = 7.8 $p < .0001$). For females, there was a significant difference between the image they believed was their real self and the one they perceived as most attractive (difference = 1.8, repeated measure t = 7.7, $p < .01$). Likewise, there was a significant difference between their perceived real image and the one they thought would be rated most attractive by other females (difference = 1.8, repeated measure t = 7.4, $p < .01$) or by males (difference = 2.2, repeated measure t = 9.3, $p < .01$). In general, females rated both their ideal figure and most attractive figure as approximately 20% smaller than their *perceived* real self and about 15% smaller than their true real self. None of these same comparisons were significantly different for male subjects.

All reported correlations are Pearson correlations. There was a significant correlation between body-image dissatisfaction and body-image distortion for females ($r = .59$, $p < .01$), but not for male participants ($r = .40$, $p > .05$).

Fear of Negative Evaluation (FNE)

There were no significant differences in mean scores between males ($M = 14.7$, $SD = 6.8$) and females ($M = 14.5$, $SD = 7.9$, independent group $t = .13$) on the FNE. There were no significant correlations between FNE and body-image distortion, or body weight dissatisfaction for either men or women. However, there was a significant correlation between FNE and body-image dissatisfaction for females ($r = .33$, $p < .01$), but not for males ($r = .436$, $p < .05$).

Body Image Avoidance Questionnaire (BIAQ)

There were no significant differences in mean scores between males ($M = 23.7$, $SD = 10.3$) and females ($M = 25.7$, $SD = 8.3$; $t = -.88$; $p > .05$) on the BIAQ. There was a significant correlation between the BIAQ and body weight dissatisfaction for females ($r = .40$, $p < .01$), but not males ($r = .57$, $p < .05$). There was also a significant correlation between the FNE and BIAQ for females ($r = .35$, $p < .01$), but not males ($r = .15$, $p > .05$).

Multidimensional Body-self Relations Questionnaire (MBSRQ)

There were no significant gender differences on any of the subscales. The results of correlations between the MBSRQ and selected variables are shown in Table 1. As can be seen, there were major differences between males and females on many of the subscales Of note are the significant correlations between the appearance evaluation scale and body weight dissatisfaction, body-image distortion, and body image dissatisfaction for females, but not for males. Similar gender differences are seem for the body areas satisfaction and weight preoccupation subscales.

Eating Disorder Inventory-II (drive for thinness and body dissatisfaction subscales only-EDI-II)

A preliminary investigation found no relevant differences between the drive for thinness and body dissatisfaction subscales when results were analyzed for each subscale separately so these two subscales were combined into one 16 question scale. Although a gender difference was found between this combined subscale (note that lower scores indicate more pathology), using a .01 level of significance, this was deemed to be nonsignificant, but this finding deserves more attention. The mean score for males was 58.8 ($SD = 16.7$) as compared to 50.8 ($SD = 14.7$) for females (independent $t = 2.1$, $p < .05$). No significant correlation was found between this scale and the FNE for either females ($r = -.437$,

TABLE 1 Correlations Between the MBSRQ Subscales[1] and Selected Variables

Measure	Males	Females
Appearance evaluation (Ae) and body weight dissatisfaction	-.37	-.48**
Ae and body image distortion	.07	-.37**
Ae and body image dissatisfaction	-.4	-.56**
Ae and FNE	-.61	-.5**
Ae and BIAQ	-.46	-.57**
Appearance orientation (Ao) and body weight dissatisfaction	.33	-.08
Ao and body image distortion	.26	.15
Ao and body image dissatisfaction	.16	.11
Ao and FNE	.19	.45**
Ao and BIAQ	.24	.23
fitness evaluation (Fe) and body weight dissatisfaction	-.25	-.18
Fe and body image distortion	.33	-.21
Fe and body image dissatisfaction	-.11	-.22
Fe and FNE	-.28	.09
Fe and BIAQ	-.51	.28
fitness orientation (Fo) and body weight dissatisfaction	-.26	-.40**
Fo and body image distortion	.24	-.16
Fo and body image dissatisfaction	-.15	-.06
Fo and FNE	-.3	-.05
Fo and BIAQ	-.44	.05
body areas satisfaction (BAS) and body weight dissatisfaction	-.49	-.35**
BAS and body image distortion	.02	-.39**
BAS and body image dissatisfaction	-.42	-.49**
BAS and FNE	-.46	-.43**
BAS and BIAQ	-.52	-.52**
subjective weight (Sw) and body weight dissatisfaction	.78**	.64**
Sw and body image distortion	.12	.30*
Sw and body image dissatisfaction	.50	.51**
Sw and FNE	.28	.06
Sw and BIAQ	.52	.31
weight preoccupation (Wp) and body weight dissatisfaction	.67**	.38**
Wp and body image distortion	.16	.30
Wp and body image dissatisfaction	.37	.55**
Wp and FNE	.16	.27*
Wp and BIAQ	.42	.62**

[1] Only relevant subscales were used. See Methods section.

** Statistically significant at the .01 level

$p < .05$), or males ($r = -.39, p > .05$). A significant correlation was found between this scale and the BIAQ for both males ($r = -.52, p = .01$) and females ($r = -.71, p < .01$). Also, there was a significant correlation between this scale and body weight dissatisfaction for both males ($r = -.50, p < .01$) and females ($r = -.50, p < .01$). A significant correlation was found between this scale and body-image dissatisfaction for females ($r = -.59, p < .01$), but not males ($r = .50, p < .05$).

DISCUSSION

Virtually all of the results of this study are consistent with previous body-image research (Rozin & Fallon, 1988; Cullari, et al., 1998). As such, this study suggests that digitally distorted images of subjects are a viable measurement technique for body-image assessment. Many of the findings of previous studies were replicated in the current study, including significant gender differences in body-image perceptions, dissatisfaction and distortion (Rozin & Fallon, 1985; Thompson, 1996; Pinhas et al., 1999).

Digital imaging appears to be a fairly easy procedure using materials that are readily available in most research settings. In the current study it took approximately 15 minutes to develop 10 distorted images for each participant. However, none of them chose any images that were more than 30% larger or smaller than their real photographs. Thus it appears that future studies can use a seven-image comparison scale with 3 images on either side of the subject's real photograph, which will further reduce preparation and research session time.

The current technique seems to solve some of the problems associated with the use of silhouettes and will probably prove to be much more effective than silhouettes with younger subjects. For example, it is much more realistic for subjects to compare themselves to an actual photograph of themselves than it is to a figure drawing. Also, this technique allows the size of the distortion among all of the images to be the same and for body proportions to remain constant. Perhaps the most significant advantage of this method is that it allows the magnitude of body-image distortion (and other differences in perception) to be more easily and accurately measured.

One of the weaknesses of this study was that the reliability of body-image assessment using digital imagining was not formally assessed. Previous studies (e.g., Thompson et al., 1994) found the reliability of the traditional silhouette procedure to be acceptable and the same is assumed for the current technique since it is very similar. In fact it offers some improvements. The first five participants in the current study were retested with the digital imaging technique after a one-week period and found to have virtually identical ratings. For this reason, this aspect of the procedure was eliminated in order to save time, however more formal reliability studies with larger samples are needed.

It appears that the most important gender differences found in this study are those of body image and body weight dissatisfaction between males and females. In virtually all the comparisons, women were significantly more dissatisfied with their own body weights and body image than males. This is despite the fact that the males in this study were about 15 pounds overweight, while the women were not. In addition, the ideal body weights and images for females, and ones they thought would be most attractive were significantly smaller than their actual ones. In males, their real, ideal and most attractive weights and body images were virtually the same.

The similarities of results between males and females found in this study are almost as interesting as the gender differences. For instance, there were no gender differences on mean scores for the FNE, MBSRQ or BIAQ scales. Similarly, the correlations between FNE and both body-image distortion and body image dissatisfaction for both males and females were similar. In addition, the ideal body-images for both males and females are approximately the same as portrayed by the popular media. For example, the ideal body-image for the females in this study was approximately 15% less than their real self. In turn, this is similar to the image of female fashion models, whose weight tends to be about 15% less than the expected weight for their height..

The results of this study suggest a bio-psycho-cultural model of body-image distortion, and perhaps eating disorders that begin with an ideal body propagated by the mass media. In females, this virtually unreachable ideal body-image results in body-image dissatisfaction, which previous studies found precedes body-image distortion (see Cullari, et al., 1998). This may set the stage for eating disorders in persons who are predisposed or are at high risk for their occurrence. If so, this would underscore the need for a prevention program for eating disorders that begins with a dramatic change in the way the mass media promotes attractiveness and sexuality in women. Although plausible, this model needs further research development.

REFERENCES

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., Text Revision). (Washington, DC: Author).

Anderson, M. (1988). *Thinking about women: Sociological perspectives on sex & gender*. NY: Macmillan.

Byrne, N. M., & Hills, A. P. (1996). Should body-image scales designed for adults be used with adolescents? *Perceptual and Motor Skills, 82*, 747-753.

Cash, T. F., & Pruzinsky, T. (Eds.). (1990). *Body images: Development, deviance and change*. NY: Guilford.

Cullari, S.,& Trubilla, R. S. (1989). Body-image distortion in normal-weight college women. *Perceptual and Motor Skills, 68*, 1195-1198.

Cullari, S., Rohrer, J. M., & Bahm, C. (1998). Body-image perceptions across sex and agegroups. *Perceptual and Motor Skills, 87*, 839-847.

Fallon, A. E., & Rozin, P. (1985). Sex differences in perception of desirable body shape*Journal of Abnormal Psychology, 94*, 102-105.

Fallon, A. E., & Rozin, P. (1988). Body image attitudes to weight, and misperceptions of figure preferences of the opposite sex: A comparison of men and women in two generations. *Journal of Abnormal Psychology, 947* 342-345.

Folk, L., Pedersen, J., & Cullari, S. (1993). Body satisfaction and self-concept of third and sixth-grade students. *Perceptual and Motor Skills, 76*, 547-553.

Garner, D. M. (1991). Eating Disorder Inventory—2 manual. Odessa, FL: Psychological Assessment Resources.

Garner, D. M. (1997). Psychoeducational principles in treatment. In D. M. Garner, & P. E. Garfinkel (Eds). *Handbook of treatment for eating disorders* (2nd ed.). (pp.147-177). NY: Guilford Press.

Gardner, R. M., Friedman, B. N., & Jackson, N. A. (1998). Methodological concerns when using silhouettes to measure body image. *Perceptual & Motor Skills, 86*(2) 387-395.

Gleghorn, A. A., Penner, L. A., Powers, P. S., & Schulman, R. (1987). The psychometric properties of several measures of body-image. *Journal of Psychopathology and Behavioral Assessment, 9*, 203-218.

Heatherton, T. F., Mahamedi, F., Striepe, M. Field, A. E., & Keel, P. (1997). A 10-year longitudinal study of body weight, dieting, and eating disorder symptoms. *Journal of Abnormal Psychology, 106*, 117-125.

Heinberg, L. J., & Thompson, J. K. (1995). Body image and televised images of thinness and attractiveness: A controlled laboratory investigation. *Journal of Social and Clinical Psychology, 14*(4), 325-338.

Kalodner, C. R. (1997). Media influences on male and female non-eating-disorder college students. A significant issue. *Eating Disorders, 5*(1), 47-57.

Metropolitan Life Insurance Company Height and Weight Table (1999). Retrieved from http://www.metlife.com/Lifeadvice/Tools/Heightnweight/.

Ogden, J., & Mundray, (1996). The effect of the media on body satisfaction: The role of gender and size. *European Eating Disorders Review, 4*(3), 171-182.

Pinhas, L., Toner, B. B., Ali, A., Garfinkel, P.E.,, & Stuckless, N. (1999). The effects of the ideal of female beauty on mood and body satisfaction. *International Journal of Eating Disorders, 25*(2), 223-226.

Rabak-Wagener, J., Eickhoff-Shemek, J., & Kelly-Vance, L. (1998). The effect of media analysis on attitudes and behaviors regarding body image among college students. *Journal of American College Health, 47*(1), 29-35.

Rauste-von Wright, M. (1989). Body image satisfaction in adolescent girls and boys: a longitudinal study. *Journal of Youth and Adolescence, 18*, 71-83.

Rosen, J. C., Srebnik, D., Saltzberg, E., & Wendt, S. (1991). Development of a body image avoidance questionnaire. *Psychological Assessment, 3*, 32-37.

Thompson, J. K. (Ed.) (1996). *Body image, eating disorders, and obesity*. Washington, DC: American Psychological Association.

Thompson, J. K. Altabe, M., Johnson, S., & Stormer, S. M. (1994). Factor analysis of multiple measures of body image disturbance: Are we all measuring the same construct? *International Journal of Eating Disorders, 16*(3), 311-315.

Stunkard, A. J., Sorenson, T., & Schlusinger, E. (1983). Use of the Danish Adoption Register for the study of obesity and thinness. In S. Kety (Ed.). *The genetics of neurological and psychiatric disorders* (pp. 115-120). NY: Raven.

Watson, D., & Friend, R. (1969). Measurement of social-evaluative anxiety. *Journal of Consulting and Clinical Psychology, 33*, 448-457.

Wegner, B. S., Hartmann, A. M., & Geist, C. R. (2000). Effect of exposure to photographs of thin models on self-consciousness in female college students. *Psychological Reports*, *86*(3), 1149-1154.

Personality and the Aesthetics of Composition: A Study of Mondrian & Hirst

Adrian Furnham and Sumedh Rao
University College London

This study found individuals were able to differentiate between originals and facsimiles of abstract art, and show preference for original works of two different artists, but that this preference was not predicted by personality variables. In all, 129 participants made preference judgements when viewing computer generated slides in a paired comparison task, choosing between original and altered works by Piet Mondrian and Damien Hirst. This was followed by a Meier Art Judgement test, the Barron-Welsh Creativity Scale, and a 'Big Five' personality questionnaire. As predicted, participants significantly preferred the original Mondrians and Hirsts as opposed to the facsimiles, supporting previous findings. Conscientiousness was found to be the only consistent significant predictor of the preference ratings.

This study has three aims: first to examine personality predictors of abstract art to see if previous findings could be replicated on the art of two famous modern abstract painters (Furnham & Walker, 2001a,b). Second, to see if findings concerning preference for original versus altered works by Mondrian can be replicated and extended to the work of Hirst; thirdly, to see if art ability, creativity or personality play any part in predicting the above preference.

One of the obstacles facing psychologists when studying aesthetic preference for the composition of paintings has been constructing aesthetically convincing and, at the same time, empirically manipulable experimental stimuli. To counter this problem, several studies have investigated aesthetic preference for very simple stimuli such as colours (McManus, Jones, & Cottrell 1981) or rectangles and triangles (McManus 1980). McManus, Edmondson, and Rodger (1985), looked at the effects on the overall "balance" of a picture by removing portions of it. McManus and Kitson (1995) investigated randomly placing dots within a frame under geometrical constraints. Locher, et al. (1999, 2001) have demonstrated that aesthetic judgement of paintings is a function partly of presentation format (gallery, slide, computer generated images).

Author info: Correspondence should be sent to: Adrian Furnham, D.Phil, D.Sc. University College London, 26 Bedford Way, London WC1 0AP. Email: a.furnham@ucl.ac.uk

North American Journal of Psychology, 2002, Vol. 4, No. 2, 233-242.
© NAJP

However none of these studies actually investigated the effect of manipulating specifically the geometry of institutionally respected and relatively well known works of art, until a study by McManus, Cheema & Stoker (1993) who used paintings by the artist Piet Mondrian as experimental stimuli. McManus et al. (1993) produced computer-generated facsimiles in which the distance between the compositional lines was altered. These facsimiles were paired with the originals and presented as slides to participants who preferred the originals significantly more than chance alone would expect. This suggested that there was something with the composition, arrangement of the lines and colours in Mondrian's paintings which constitute their specific artistic and aesthetic appeal.

Noll (1966) conducted a study along similar lines to the McManus et al. (1993) study using pairs of original and 'fake' pictures from the same artist. Results showed that when paired with the originals, participants significantly preferred the computer generated "alternative" version. This study aimed to replicate McManus et al.'s (1993) study on a different group of participants to see if their findings are robust. In addition this study investigated whether this effect is only found with paintings by Piet Mondrian or whether this effect is found with other artists' work such as Damien Hirst. Damien Hirst (1965-) is a modern artist who works in sculpture and paints.

This study manipulated Hirsts in a similar way to the manipulation of Mondrians in the McManus et al. (1993) study. The size of the grid in the original pharmaceutical paintings, in which the dots are placed, was increased or reduced by 50% while the dots' size and colour was kept the same. The grid size was manipulated as changing the colours of the dots does not affect the composition of the painting but the 'mood'. Hirst himself stated how, though the colours seem random, the colours he used were affected by his mood, to produce different paintings with different emotions. Changing dot size would have the same effect as changing grid size – bigger space between the dots in comparison to the size of the dots themselves.

Thus it was possible to test the same idea on a different type of picture namely that people would prefer originals which we carefully and thoughtfully balanced compared to altered spatial arrangements in the picture.

Hypothesis 1. Participants will show significant preference for the original Mondrians and the original Hirsts more than the facsimiles. The second feature of this study was to examine personality trait correlates of painting preference. One of the first investigations into preference in art was by Eysenck in 1940. There have been several studies trying to determine how different personality variables affect preference for different styles of art (Furnham & Bunyan, 1988; Furnham & Walker, 2001ab; Rawlings, Vidal & Furnham, 2000; Toback, Myers, & Bailey, 1981; Wilson, Ausman, & Matthews, 1973). Furnham and Walker (2001ab) found that Neuroticism and Openness (to experience) were correlated with preference for abstract paintings; Neuroticism, Openness, and Agreeableness were correlated with pop-art; and Openness and Conscientiousness were correlated with representational art. In addition sex, art level studied, history of art level studied, and frequency of visits to art galleries were correlated with abstract art preference.

Hypothesis 2. Preference for the originals will be correlated with Neuroticism, Openness to experience, sex, art level studied, and history of art level studied. This prediction is based on the hypothesis that the findings of preference relating to abstract art found by Furnham and Walker (2001ab) will map directly onto the preference task investigating Mondrians and Hirsts. Consequently those variables which contribute most strongly to abstract art (e.g. Neuroticism, Openness to experience, art level studied) will most influence preference for the originals by allowing greater sensitivity to the abstract art of Mondrian and Hirst. Those variables which are not related to abstract art preference (e.g. Conscientiousness, Extraversion) will not be related to this art task.

Further it is hypothesised that participants individually motivated to visit art galleries, as well as those who enjoy visiting galleries sufficiently to go by themselves is related to greater liking of art and "sophistication in art" resulting in preference for the originals. Various studies have demonstrated that knowledge of art are closely related to aesthetic judgements (Furnham & Chamorro-Premuzic, 2002; Smith & Wolf, 1996).

Hypothesis 3. Frequency of Art gallery visits will correlate with preference for originals. In addition to investigating the 'Big Five' personality variable and background variable measures two other scales will be investigated: the *Meier Art Judgement test*, measuring aesthetic appreciation, and the *Barron-Welsh* test, measuring creativity. The *Meier Art Judgement test* was published by Meier and Seashore in 1929 and later updated (Meier, 1940). It consists of pairs of black and white pictures that differ in only one feature (e.g. a moved tree, or hidden sun). The majority of the research with the original test dated from 1941 and before. In 1940 the test was revised and renamed the *Meier Art Judgement* test. The *Meier Art Judgement* test has also been shown to be related to a liking for representational art (Hill & Junus, 1979).

Clearly being able to differentiate between originals and facsimiles maybe seen as a validation of any art judgement test.

Hypothesis 4. The *Meier Art Judgement* scale score will positively correlate with preference for the original Mondrians and Hirsts.

There are few psychometrized tests of creativity. The *Barron-Welsh* scale (Welsh, 1975) is a preference scale for several monochrome drawings varying in complexity, with ratings, like to dislike, given for the complex or simple drawings. Ratings on the *Barron-Welsh* scale have been identified in several studies as being a good indicator of creativity (Eysenck & Furnham 1993; Ridley 1981, Schmidt, 1974).

Hypothesis 5. The *Barron-Welsh* scale will be positively correlated with preference for original Mondrians and Hirsts.

METHOD

Participants

129 participants took part in this study, comprised of pupils from a sixth-form college in Winchester. There were 52 males and 77 females, ranging in age from 16 to 19 years. The mean age was 16.88 years (*SD*=.69). They took part in the study as part of their Psychology A-level course and were not paid.

Materials

Materials used were a 35mm slide projector, 100 stimuli slides, and questionnaires.

Mondrians The study used part of the stimuli slides from the McManus, Cheema and Stoker (1993) study. Several Mondrian paintings from his mature period were used which consisted of a background of pure white with black vertical and horizontal lines, extending either to the frame, or to another black line and with solid areas of primary colours between the lines. Each line therefore can be expressed as a proportional distance between two other lines. These proportional distances were adjusted according to a series of logistic transformations. This method allowed proportions at the middle of the range and at the extremes to be altered by a proportionately similar amount, and therefore without changing the fundamental syntax of the picture.

Hirsts In a similar manner to Mondrian, 25 paintings of Damien Hirst's 'Pharmaceutical collection' were scanned using a colour scanner. These pictures were then adjusted using the graphics package Adobe Photoshop 5.0 on an IBM PC. Facsimiles of Hirsts were produced by altering the distance between the dots by 50% by cutting and pasting. With the pseudo-Hirsts half the slides had the distance increased by 50%, and the other half reduced by 50%, and half were placed on the left of the original and half on the right. With the facsimiles the dots remained the same size and colour throughout, only the distance between the dots was varied. This resulted in 22 test slides and 3 practice slides.

Meier Art Test The original test consists of 100 pairs of pictures resembling etchings, one being the original and one that differs in a certain feature which the participant is made aware of. Of these 25 were chosen with equal number of correct responses being right and left, to prevent bias towards one side. This scale has shown acceptable levels of validity (Hevner, 1933) and an alpha-coefficient of .71 found prior to analysis.

Barron-Welsh This test consists of 86 pictures which appear like sketches with differing complexity. 24 of these form the 'Like' scale, 38 of these form the 'Dislike' scale, and the rest are fillers. Based on this original ratio of Like to Dislike, 9 Like pictures and 13 Dislike pictures were produced onto 35mm slides.

Questionnaire The questionnaire consisted of 3 sections:

1. Background variables Sex and age, ethnic origin, as well as the level of education.

2. Six specific questions: How much have you studied art? How much have you studied history of art? How often have you visited art galleries this year? How often do you intend to visit art galleries in the next year? When you did go to art galleries, who did you go with on the whole? On what proportion of visits did you actively choose to go?

3. NEO Five-Factor Inventory (FFI) Form S This questionnaire measures the "Big five" factors of personality: Openness to experience, Conscientiousness, Extraversion, Agreeableness and Neuroticism. It contains 60 self-descriptive statements - 12 for each factor.

Procedure

The study was conducted in a small lecture hall in groups of 15 to 20. Time was allowed to complete the background questions then the slides were presented in the order Mondrians, Hirsts, *Meier Art Judgement Test*, and then *Barron-Welsh*. Each slide was presented for 15 seconds. After presentation of the slides participants were asked to complete the final NEO-FFI section. Participants were questioned afterwards as to whether they were familiar with either Mondrian or Hirst. There was no reported difference.

RESULTS

1. Alpha-Reliability: The *Meier Art Judgement* and *Barron-Welsh* scales showed sufficient internal reliability with alphas of 0.66 and 0.86 respectively. The NEO-FFI also had adequate reliability with alpha scores for Neuroticism 0.85, Extraversion 0.73, Openness 0.73, Agreeableness 0.74 and Conscientiousness 0.71.

However, the Hirst scale and especially the Mondrian scale showed very poor alpha scores of 0.54 and 0.16 respectively. Based on this, items in these originally 22-item scales were removed until reasonable alphas were found. 4 items were removed from the Hirst scale to produce an acceptable score, but 9 items had to be removed to improve the score, whereupon further item removal failed to add sufficiently to alpha scores. The final scales were: Mondrian (13-item scale): *Alpha* = *0.53;* Hirst (18-item scale): *Alpha* = *0.71*

2. Judgement scores: The 6-point scale was collapsed into a 2-point scale indicating whether or not the subject had preferred overall the original or the facsimile, giving total numbers of "correct" judgements for each participant.

First the analysis was done on the 2-point scale with the 13-item Mondrian scale using 123 participants, of 1,599 judgements, 917 (57.4%) were correct. Participants were correct in an average of 7.46 cases out of 13 (SD=2.23), which is significantly higher than chance expectation of 6.50 (t=9.73, p<.01). Second the analysis was done on the 6-point scale, responses were awarded between 0 (for much preferring the incorrect response) to 5 (for much preferring the correct response). Participants scored an average of 34.51 cases out of 65 (SD=5.42), which is significantly higher than chance expectation of 32.5 (t=4.12, p<.01). With the 18-item Hirst scale and using the two point scale on 128 participants, of 2,304 judgements, 1,231 (53.4%) were correct. Participants were correct in an average of 9.61 cases out of 18 (SD=3.22), which is significantly higher than chance expectation of 9 (t=2.15, p<.05). A similar analysis was repeated on the six-point scale. Participants scored an average of 45.63 cases out of 90 (SD=9.08), which is not significantly higher than chance expectation of 45 (t=0.75, not significant). Thus using the two-point preference scale there is support for hypothesis 1.

3. *Correlations:* Preferences on the six-point scale were correlated (using Pearson's r) with the 16 personality measures and background variables resulting in a total of 32 correlations. There were more significant correlations than were expected by chance (3 correlations out of 32, with 1.6 expected by chance with p<.05). Mondrian (the 13-item scale) preferences was found to be correlated with Conscientiousness (r = 0.21, p<.05). Hirst (18-item scale) also correlated significantly with Conscientiousness (r= 0.29, p<.01) and with Sex (r = 0.24, p<.01). There was therefore no support for Hypothesis 2 as no other correlations were significant.

4. *Multiple Regressions*

Mondrians: Several block-entry multiple regressions were carried out with the 13-item Mondrian scale and various combinations of personality and background variables. None produced significant multiple regressions. The only significant regression was a linear regression with conscientiousness as the predictor (beta=0.22 t=2.44 adj r^2=0.04). ($F_{(1,118)}$ = 5.94, p<.05).

Hirsts: The overall regression accounted for 7.4% of the variance. Conscientiousness was the only variable found to be significant. The correlation between sex and Hirst score failed to show sex as significant when entered as a variable. This can be explained by sex accounting for some of the variance of Conscientiousness, which can be tested by regressing Sex on Conscientiousness. This regression found sex to be a significant predictor of Conscientiousness (beta=0.28, t=3.24, adj r^2=0.07), ($F_{1,124}$ = 10.46, $p<.01$). The variables found relating to preference tasks: Neuroticism, Openness to experience, sex, art level studied, history of art level studied and frequency of visiting art galleries, were not related to preference for originals. In addition, Conscientiousness *was* found to be a predictor as opposed to the prediction. This shows there was no empirical support for hypotheses 2, 3, 4 and 5.

5. Meier Art Judgement Test With the full 22-item scale using 129 participants, of 2,794 judgements, 2,038 (72.9%) were correct. On the 2-point scale participants were correct in an average of 16.04 cases out of 22 (SD=2.61), which is significantly higher than chance expectation of 11 (t=9.73, $p<.01$). On the 6-point scale participants scored an average of 70.80 cases out of 110 (SD=9.41), which is significantly higher than chance expectation of 55 (t=18.91, $p<.01$).

Preferences on the six-point scale were correlated (using Pearson's *r)* with the other personality measures and background variables resulting in 16 correlations. Again there were more significant correlations than were expected by chance (2 correlations out of 16 with 0.8 expected by chance with $p<.05$). The *Meier* scale correlated significantly with Conscientiousness (r = 0.21, $p<.05$) and with sex (r = 0.18, $p<.05$). A block entry multiple regression was found to be significant with all variables (including ethnic origin). To show how the variance changes through adding variables, blocks of different variables were entered in sequence: firstly *Barron-Welsh*, then the 'Big 5', then demographic variables, and finally art questions.

The first regression with just the *Barron-Welsh* variables was found to be nonsignificant. Entering the Big 5 to the existing *Barron-Welsh* equation made the overall regression significant ($F_{(6,\ 113)}$ = 2.20, $p<.05$; R^2_{adj} = .05) and found Conscientiousness to be significant (beta = .24). Adding demographic variables adds to the variance accounted for from 6% to 16% ($F_{9,\ 110}$ = 2.26, $p<.05$; R^2_{adj} = .16), but reduces the significance of Conscientiousness (beta = .20; from $p<.12$ to $p<.36$), indicating that sex is accounting for part of the effect. Finally regressing all variables onto Meier *Art Judgement* reduces the variance accounted for to 13% ($F_{15,\ 93}$ = 2.05, $p<.05$; R^2_{adj} = .13). In addition, Openness to Experience was found to be significant in this equation once the art questions were added. Openness to Experience (beta = .27) and Conscientiousness (beta

= .20) were found to be significant predictors in the final equation. The emergence of Openness to experience as a predictor in the last equation suggests that the art background variables have some interaction effects with Openness resulting in the variable's significance.

DISCUSSION

This study found that participants preferred the original Mondrian paintings at levels significantly greater than chance, indicating that the McManus *et al.* (1993) findings are robust. However, it seemed that participants had greater difficulty in differentiating the Hirsts compared to the Mondrians. This conclusion is reflected by the percentage of correct judgements (57.4% correct judgements with the Mondrians, whereas only 53.4% correct with the Hirsts). This may be due to the fact that in this study we manipulated the size of the dots rather than their colour which is a rather different form of variation.

This study failed to support hypothesis 2 and 3 about relationships between preference for originals and personality, background and art experience variables. Hypotheses 4 and 5, proposing correlations with *Meier* and *Barron-Welsh* were also not supported. The aesthetic judgement evaluated with *Meier* test was not correlated with aesthetic judgements of Mondrians or Hirsts. This may be due to either the *Meier* not being a valid measure of aesthetic judgement in general or that this scale which measures aesthetic judgement with representational art can not be applied with abstract art judgement. In addition it was found that creativity as measured by the *Barron-Welsh* scale was not related to preference for original Mondrians or Hirsts, which would suggest that ability to create art does not influence sensitivity to spatial composition of abstract art.

Studies requiring participants to rate pictures of different styles (e.g. Wilson et al, 1979; Tobacyk et al, 1983) found that certain personality traits are linked to art preference, but in this specific choice task involving art they are not systematically related. An individual choosing between an original and a facsimile of a work of modern art may invoke a different process to when an individual rates a piece of art on a preference scale. With preference judgements social factors are integral, whereas with a paired comparison task cognitive processes are more important. The latter have more of a good-bad quality whereas the former are more often associated with liking. It is as if the former is a power test and the latter is a preference test.

Whereas a personality trait may contribute to an individual liking abstract art, it does not improve the ability of an individual to correctly identify the original paintings. An individual who shows a preference for abstract art (which would be indicated by high scores on certain personality traits) would not be better off at recognising an original work of abstract art, in a blind paired comparison experiment.

Regressions on the *Meier Art Judgement* test found Openness to Experience and Conscientiousness to be significant predictors. The *Meier* scale scores were not to be related to correctly identifying either Mondrians or Hirsts, showed a relationship with other variables. Conscientiousness again showed a significant relation suggesting that in paired comparison tasks this variable may be an important factor in correctly identifying the originals. The *Barron-Welsh* creativity score did not show a significant relationship with *Meier* scale indicating that *Meier* scale may not be a good predictor of creativity, which was the original purpose of the scale.

The *Meier* scale has been found to be a mediocre predictor of judgement for Mondrians and Hirsts and therefore not a good predictor of aesthetic judgement, at least on this sort of task. The findings do not counter the criticisms against it (Hill & Junus, 1979; Karaeng & Sandstroem, 1959) despite participants choosing the correct option significantly more than chance. The prediction that a bipolar 'good taste' factor would be found which affects judgement scores has not been confirmed.

This modest study did have limitations. It had a relative small sample of participants and may have profited by exploring their art knowledge or interest in greater depth. Second, the quality and size of the slides could have affected ratings and made them more difficult to distinguish (Locher et al, 2001). Thirdly, there remains some doubt about the psychometric validity of both the art judgement and creativity test, which may have accounted for the more significant findings. However, the psychology of aesthetics is a small but growing field that is not only of interest to academic psychologists but art educators, gallery professionals and those more generally in the art business as it helps them target certain groups.

REFERENCES

Eysenck H. J. (1940). The general factor in aesthetic judgements. *British Journal of Psychology, 31,* 94-102.

Eysenck H. J., & Furnham, A. (1993). Personality and the Barron-Welsh Art Scale, *Perceptual and Motor Skills, 76,* 837-838.

Furnham, A., & Bunyan, M. (1988). Personality and art preferences. *European Journal of Personality, 2,* 67-74.

Furnham, A., & Chamorro-Premuzic.T. (2002). Personality, intelligence and interest in art. Paper under review.

Furnham, A. & Walker J. (2001a). Personality and preference for abstract, pop art and representational paintings. *European Journal of Personality, 15,* 57-72.

Furnham, A. & Walker J. (2001b). The influence of personality traits, previous experience of art, and demographic variables on artistic preference. *Personality and Individual Differences, 31,* 997-1017.

Hill, K.A. & Junus, F (1979). Individual Differences in Concept Learning of Painting styles. *Perceptual and Motor Skills, 49,* 255-261.

Hevner, K. (1933) A method of correcting for guessing in true-false test and empirical evidence in support of it. *Journal of Social Psychology 3, 359-362.*

Locher, P., Smith, L.F., & Smith, J.K. (1999). Original paintings versus slide and computer reproductions: A comparison of viewer responses. *Empirical Studies in the Arts, 17,* 121-129.

Locher, P., Smith, J.K., & Smith, L.F. (2001). The influence of presentation format and viewer training in the visual arts on the perception of pictorial and aesthetic qualities of paintings. *Perception, 30,* 449-465.

McManus, I.C. (1980) The aesthetics of simple figures. *British Journal of Psychology 71,* 505-524.

McManus, I.C., Cheema, B., & Stoker J. (1993) The Aesthetics of Composition: A Study of Mondrian, *Empirical Studies of the Arts 11,* 83-94.

McManus, I.C., Jones, A.L., & Cottrell, J. (1981). The aesthetics of colour. *Perception, 10,* 651-666.

McManus, I.C., Edmondson, D., & Rodger, J. (1985). Balance in pictures. *British Journal of Psychology, 76,* 311-324.

McManus, I.C. & Kitson, C. (1995) Compositional geometry in pictures. *Empirical Studies of the Arts 13(1),* 73-94.

Meier, N.C., & Seashore, C.E. (1929) *The Meier-Seashore Art judgement test.* Iowa City, Bureau of Educational Research, University of Iowa.

Meier, N.C. (1940) *The Meier art tests I. Art Judgement* Iowa City, Bureau of Educational Research, University of Iowa.

Noll, A.M. (1966) Human or Machine: A Subjective Comparison of Piet Mondrian's Composition with lines (1917) and a Computer-Generates Picture, *Psychological Record, 16,* 1-10.

Rawlings, D., Vidal, N., Furnham, A. (2000). Personality and aesthetic preference in Spain and England: Two studies relating sensation seeking and openness to experience to liking for paintings and music. *European Journal of Personality, 14,* 553-576.

Ridley, D.R. (1981) Barron-Welsh scores and creativity: A second look. *Perceptual and Motor Skills 49(3),* 756-758.

Schmidt, H.E. (1974) The identification of high and low creativity in architecture students. *Psychologia Africana 15,* 15-40.

Smith, J.K., & Wolf, L.F. (1996). Museum visitor preferences and intentions in constructing aesthetic experience. *Poetics: A Journal of Empirical Research in the Arts and Literature, 24,* 219-230.

Tobacyk, J.J., Bailey, L, & Myers, H. (1979). Preference for paintings and personality traits. *Psychological Reports, 45,* 787-793.

Tobacyk, J.J., Myers, H., & Bailey, L. (1981) Field-dependence, sensation seeking and preference for paintings. *Journal of Personality Assessment, 45,* 270-277.

Welsh, G. (1975). *Creativity and Intelligence: A Personality Approach.* Chapel Hill, NC: Institute for Research in Social Science.

Wilson, G.D., Ausman, J. & Mathews, T.R. (1973). Conservatism and art preferences, *Journal of Personality and Social Psychology, 25,* 286-289.

A Short Scale of Family Atmosphere (SOFA): Development and Psychometric Evaluation

Geoffrey N. Molloy
Monash University, Melbourne, Australia

Julie F. Pallant
Swinburne University of Technology, Melbourne, Australia

A newly developed measure of family harmony, the *Scale of Family Atmosphere* (SOFA), is described and evaluated. The brief 10-item questionnaire is time-efficient and uses language that is readily understood by respondents ranging in age from early to late adolescence. The SOFA was administered, with a battery of other scales, to a sample of 224 adolescents (106 males, 118 females), ranging in age from 12yrs to 18yrs (*M*=15yrs, *SD*=2.2yrs). Analyses of key psychometric properties of the SOFA indicate that the scale has good internal consistency (Cronbach alpha=.87), with all items loading on a single underlying factor. The validity of the scale is supported by its expected associations with measures of psychological wellbeing (anxiety, depression, self esteem), and the personality dispositions of neuroticism and psychoticism. Scores were found to be relatively independent of verbal ability and socially desirable response bias.

Numerous studies have reported statistically and socially significant links between adolescents' self-reports of family atmosphere and psychological adjustment. One of the most widely used measures of domestic atmosphere is the family environment scale (FES; Moos & Moos, 1986). The FES is one of ten social climate scales developed by Rudolf Moos and his colleagues and comprises 90 true-false items purporting to assess three dimensions of family environment described as relationships, personal growth and system maintenance. Ten subscales make up the three dimensions.

Despite its popularity, several surveys evaluating the FES have questioned its reliability and validity and thus the scale's usefulness as a research instrument (Boyd, Gullone, Needleman & Burt, 1997; Roosa & Beals, 1990a, b). Roosa and Beals summarised responses of 385 families and reported that five of the ten FES subscales yielded Cronbach alpha internal consistency values of less than .70. More recently Boyd and her colleagues (1997) administered the FES to 1,289 Australian adolescents.

Author info: Correspondence should be sent to: Dr Julie F. Pallant, School of Mathematical Sciences, Swinburne University of Technology, PO Box 218, Hawthorn. VIC. 3122. Australia. Email: jpallant@swin.edu.au.
North American Journal of Psychology, 2002, Vol. 4, No. 2, 243-248.
© NAJP

Alpha reliability estimates for this sample on the ten subscales ranged from .39 for *Independence* to .72 for *Conflict*. Only two of the ten subscales yielded alpha values in excess of .70. In the light of these data Boyd et al. (1997) argued that, regardless of the type of scale, reliability estimates of such magnitude are simply inadequate for research purposes (see also; Molloy, 2001; Nunnally, 1978; Pallant, 2001; Tabachnick & Fidell, 1996) and called for a revision of the scale. In answer to our colleagues' challenge, we devised a 10-item scale of family atmosphere with a sample of Australian adolescents. Our aim was to develop a global scale with items that were written in plain English, simple to administer and score (temporally economic), internally consistent and structurally sound.

METHOD

Participants

Data for this investigation were gathered as part of a class project for fourth-year psychology students enrolled in the Postgraduate Diploma of Psychology (Education) at Monash University, Melbourne, Australia. The volunteer sample consisted of 224 adolescents: 106 males (47.3%), 118 females (52.7%) aged 12 to 18 years (*M*=15.07 yrs, *SD*=2.15) attending coeducational high schools in Melbourne and rural Victoria. Ethical approval was obtained from the university's ethics committee, along with the consent of schools, participants and participants' parents or guardians.

To ensure anonymity, participants were asked to put their completed questionnaire booklets in plain A4 envelopes, seal the envelopes and place them in a collection box provided for this purpose. Testing was conducted with assistance from class teachers during normal class periods under group conditions in a one-hour period dedicated for this purpose.

Measures and Procedure

The Scale of Family Atmosphere (SOFA) was administered as part of a battery of questionnaires consisting of: the Eysenck Personality Questionnaire (EPQ; Eysenck & Eysenck, 1975), including the Extraversion, Neuroticism, Lie and Psychoticism scales; the Rosenberg Self-Esteem Scale (SES; Rosenberg 1965); a modified form of the Wechsler Similarities Subtest (Wechsler, 1992); and Reynolds' Adolescent Depression Scale (RADS; Reynolds, 1987). Except for the SOFA, the specific item content and psychometric properties of the construct validation measures are described in the references cited.

The 10 items of the SOFA were selected from a large item pool generated by a class of postgraduate psychology students as part of a class assignment. Each student was asked to write a list of 10 items that they believed tapped the broad domain of "family atmosphere." From this item pool a subset of 15 of the most commonly occurring items were identified. The correlations among the items were inspected and reliability analyses performed. Items with low item-total and inter-item correlations were removed. This resulted in the retention of 10 items (presented in the Appendix). A 5-point response scale is used where 1=strongly disagree and 5=strongly agree. (Footnote 1).

RESULTS AND DISCUSSION

The 10 items of the SOFA were subjected to a variety of statistical analyses designed to assess the scale's internal reliability, underlying factor structure, and construct validity. The scale showed good internal consistency with a Cronbach's alpha value of .87. Principal components analysis revealed two factors with eigenvalues exceeding 1, explaining 46.9% and 12.97% of the variance, respectively. Inspection of the PCA results indicate that all items loaded above .54 on the first major component extracted, suggesting a strong underlying dimension. The screeplot supported the retention of two components for further investigation. Both orthogonal (Varimax) and oblique (Oblimin) rotations were performed, yielding consistent results. Factor loadings for the 10 items following Varimax rotation are presented in Table 1.

TABLE 1 Rotated Factor Loadings of SOFA Items for Two Factor Solution Following Varimax Rotation

		Factor 1	Factor 2
Q7	I have a happy and close relationship with my father	.797	
Q4	I have a happy and close relationship with my mother	.761	
Q1	My childhood has been a happy one	.721	
Q8	My family listen and take notice of what I say	.719	
Q5	My house is full of tensions and disagreements*	.717	
Q2	I respect my parents	.678	
Q3	My parents fight a lot*		.818
Q9	My parents rarely argue		.791
Q10	Sometimes members of my family are physically violent to each other*		.738
Q6	I enjoy being around my family	.500	.539
	Eigenvalue (after rotation)	3.58	2.41
	% of variance (after rotation)	35.79	24.07

Note. Only loadings above .3 are displayed. * Item reverse scored.

Items 1, 2, 4, 5, and 8 loaded strongly (above .67) on Component 1. The marker statements for Component 2 were items 3, 9 and 10. Item 6 ('*My house is full of tension and disagreements*') showed approximately equal loadings on both components. This solution suggested some separation of the three items that tapped parental conflict (eg. '*My parents fight a lot*'). The two sets of items identified appear to loosely correspond to the Cohesion and Conflict dimensions of the FES.

Although the two component solution warrants further investigation, collectively the strong loadings of all items on the first major component extracted, the substantial correlation (r=.48) between the components, the high Cronbach's alpha value, and the small number of items loading on the second component suggest that it would be more appropriate at this stage to combine all items to form a single scale. Further investigation of the stability of the factor structure and the potential usefulness of the separate sub-factors is required, using a variety of samples.

An independent groups t-test revealed no significant sex differences between males (M=38.6, SD=6.89) and females (M=38.12, SD=7.35, $t_{(207)}=.49$, $p=.63$) on the SOFA. Scale scores were independent of age: $r_{(209)}=-.04$, $p=.60$). There was a very weak relationship $r_{(206)}=-.15$, $p=.04$) between SOFA scores and the Lie subscale of the EPQ, indicating that participants' reactions to SOFA items are not unduly influenced by socially desirable responding. The fact that SOFA scores were unrelated to a measure of verbal ability ($r_{(208)}=.05$, $p=.44$) verifies that the item content was appropriate for the age range sampled. Further support for the SOFA's construct validity was provided by its predicted links with measures of depression ($r=-.45$, $p<.0005$), anxiety ($r=-.32$, $p<.0005$), self esteem ($r=.36$, $p<.0005$), neuroticism ($r=-.23$, $p=.001$) and psychoticism ($r=-.34$, $p<.0005$). In other words, high scores on the SOFA are associated with better psychological well-being on these validation measures.

The results of this study provide preliminary support for the psychometric properties and potential usefulness of a global, but brief, measure of family atmosphere. Further investigation of the scale's factor structure, test-retest reliability and predictive validity is required, however this study provides support for the SOFA's internal consistency and construct validity in an adolescent sample. For researchers requiring a quick, easy to administer, indicator of family atmosphere, the SOFA appears to offer a promising alternative to the much longer, and more complex scales currently available.

REFERENCES

Boyd, C. P., Gullone, E., Needleman, G. L., & Burt, T. (1997). The Family Environment Scale: Reliability and normative data for an adolescent sample. *Family Process, 36*, 369-373.

Eysenck, H.J., & Eysenck, S.B.G. (1975). *Manual of the Eysenck Personality Questionnaire*. London: Hodder and Stoughton.

Molloy, G. N. (2001). *Some psychometric issues in scale development.* Paper presented at the Annual Victorian Conference of the Australian Association for Cognitive and Behavior Therapy. Australian Catholic University, Melbourne, Australia.

Moos, R. H., & Moos, B. S. (1986). *The Family Environment Scale: The Manual.* Palo Alto, CA: Consulting Psychologists Press.

Nunnally, J.O. (1978). *Psychometric theory*. New York: McGraw-Hill.

Pallant, J. F. (2001). *SPSS Survival Manual: A step by step guide to data analysis using SPSS*. Sydney: Allen and Unwin.

Roosa, M. W., & Beals, J. (1990a). Measurement issues in family assessment: The case of the Family Environment Scale. *Family Process, 29*, 191-198.

Roosa, M. W., & Beals, J. (1990b). A final comment on the case of the Family Environment Scale. *Family Process, 29*, 209-211.

Rosenberg, M. (1965). *Society and adolescent self-image*. Princeton, N.J.: Princeton University Press.

Reynolds, W. M. (1987). *Reynolds Adolescent Depression Scale manual*. Florida: Psychological Assessment Resources, Inc.

Tabachnick, B. G., & Fidell, L. S. (1996). *Using multivariate statistics* (3rd ed.). New York: Harper Collins.

Wechsler, D. (1992). *Wechsler intelligence scale for children (3rd edition) manual (Australian adaptation)*. San Antonio, Texas: Harcourt Brace Jovanovich.

Note: Copies of the SOFA are freely available for use in research and clinical settings. Requests should be sent to the second author: Dr Julie F. Pallant at jpallant@swin.edu.au.

APPENDIX

Scale of Family Atmosphere

Instructions: Here are some statements about how you, or people in your family, might feel or act. Think about each statement and tick the box that indicates how much you agree or disagree with each statement.

1 My childhood has been a happy one
2 I respect my parents
3 My parents fight a lot*
4 I have a happy and close relationship with my mother
5 My house is full of tensions and disagreements*
6 I enjoy being around my family
7 I have a happy and close relationship with my father
8 My family listen and take notice of what I say
9 My parents rarely argue
10 Sometimes members of my family are physically violent to each other*

Note: Each item is followed by five boxes. Choices are indicated below.

Strongly disagree	disagree	neither agree nor disagree	agree	strongly agree

Items marked with an asterisk (*) are reverse scored. A total score is calculated by adding scores on all items.

The Sport Widow Hypothesis: A Research Note

Gordon W. Russell and Robert L. Arms
University of Lethbridge

University females ($N = 75$) were asked if their partners' involvement in sports strengthened or weakened their romantic relationships. Involvement in sports was positively related to the perceived strength of relationships. However, their partners' involvement in sports was unrelated to a measure of relationship closeness and their love/liking for their partner. In a second study involving older, non-university couples ($N = 91$), females' reports of strength of relationship, relationship closeness and affect for partner were unrelated to their partners' involvement in sports. Males saw their relationships strengthened insofar as positive correlations were found between their partners' involvement and both the perceived strength of their relationships and relationship closeness.

By one estimate, approximately two-thirds of Americans regard themselves as sport fans (Wann, Melnick, Russell, & Pease, 2001). The enthusiasm of fans for their favorite sport(s) varies widely. Some are passive and irregular followers of their sport. Others are seemingly consumed with a passion that overshadows other life pursuits, e.g., job, education. Does this passion for sports generally strengthen or damage the quality of relationships? Speculation in the media favors the latter view and has popularized the phrase "sport widow." Here the long-suffering woman is saddled with a man who largely ignores his family and domestic responsibilities in favor of his love of sports.

Quirk (1997) ostensibly found support for a disruptive influence on relationships by those he calls "sportsaholics," i.e. almost exclusively men who are so addicted to sport that their relationships with wives or girlfriends are imperiled. However, his methodology is flawed (see Wann et al., 2001, pp. 162-163 for a critique). Additionally, Smith, Patterson, Williams, and Hogg (1981) report that among avid male sports fans, 26% acknowledge television viewing is a source of friction in their relationships.

Roloff and Solomon (1989) examined the impact of sports on relational harmony among a sample of university students, 42% of whom reported on their "lapsed" relationships. They found "... no support for the notion that conflict over sports adversely affects relational quality"

Author info: Correspondence should be sent to: Professor Gordon W. Russell, 1 McKenzie Crescent, Sidney, British Columbia, Canada V8L 5Y7. E-mail russell@uleth.ca.

North American Journal of Psychology, 2002, Vol. 4, No. 2, 249-252.

© NAJP

(p. 307). Based on telephone interviews, Gantz et al. (1995a) similarly found that watching major sports on television "... appears to be a minor and nondisruptive activity in most ongoing relationships" (p. 352; see also Gantz et al., 1995b).

The present investigation was intended to widen the nomological net, using a different cultural sample and methodology. Additionally, we extended the scope of the investigation to include an older, non-university sample of couples.

STUDY 1

Method

Participants. The participants (N = 75) were female volunteers recruited from the introductory psychology subject pool at the University of Lethbridge. The sign-up folder specified an interest in those presently involved in a romantic relationship. The mean age of the women was 20.6 years (*SD* = 4.6) and that of their partners 22.1 (*SD* = 4.9). Eleven percent were married, 4% engaged and 7% cohabiting. A further 73% were dating only this person while 5% were dating this person and others. Finally, the couples had known each other for an average of 3.6 years (*SD* = 4.7). Their participation earned a 1 % bonus in the course.

Procedure. Participants were asked to anonymously complete a battery of measures assessing the current status of their relationships with a dating partner or someone with whom they were romantically involved. The measures were administered to groups of three to six persons, a task that took approximately 20 minutes.

Measures. The battery included a measure of influence of sports involvement in weakening or strengthening a relationship and Berscheid, Snyder, & Omoto's (1989) *Relationship Closeness Inventory* (RCI) and *Affect for Partner Index*. Also, a composite index assessing the extent of their partner's involvement in sport was administered.

The participants rated the influence of their partner's interest in sports in weakening or strengthening the relationship on a 7-point scale anchored by "weaken" and "strengthen" with the mid-point labeled "neither." Values ranging from -3 through 0 to + 3 underlie the scale points. The RCI (r = .82, test-retest) is comprised of three subscales, i.e., Frequency, Diversity and Strength, that are combined to form an overall measure of relationship closeness. The *Affect for Partner Index* represents the combined responses to two items asking: "how much do you *love* X (*like* X)?" Seven-point scales anchored by "not at all" and "a great deal" accompanied each item. The extent of a partner's involvement was measured by five items assessing the weekly frequency of their participation as a spectator, television viewer, an athlete/coach/parent, reading about sports in newspapers/magazines and discussing the topic with others. Seven-point rating scales underlaid the items and were anchored by "Never" and "Extremely Often." The presentation of measures was counterbalanced to minimize order effects.

Results

Analyses of female responses yielded results consistent with earlier findings indicating a lack of support for the notion that mens' involvement in sports tends to weaken relationships (Gantz et al., 1995a, 1995b; Roloff & Solomon, 1989). As shown in Table 1, neither the RCI nor the *Affect for Partner Index* scores was related to their partners' involvement in sports. Nevertheless, the co-eds expressed the view that their relationships were strengthened by their partners' participation.

TABLE 1 Correlations Between Partners' Involvement in Sport and Relationship Measures

	Study 1	Study 2	
	Females $N = 75$	Females $N = 91$	Males $N = 91$
Relationship Strength	.25*	.10	.38***
RCI	.06	.12	.24**
Affect for Partner	.14	-.03	-.01

* $p < .05$ (two-tailed test)
** $p < .02$
*** $p < .001$

STUDY 2

The investigation was expanded in a second study to include somewhat older, non-university couples.

Participants. The participants were unpaid, volunteer couples ($N = 91$) recruited by members of a senior social-experimental class during a semester break. The sample was comprised of couples contacted by the students in their home towns and geographically represented Alberta and its neighboring provinces, British Columbia and Saskatchewan, in addition to the Northwest Territories. The females had a mean age of 29.9 ($SD = 12.1$) whereas the mean age of males was 31.5 ($SD = 13.3$). Forty eight percent were married, 6% engaged and 10% living together, while 31% were dating only this person, with a further 5% dating this person and others. The couples had known each other for an average of 9.6 years ($SD = 11.8$).

Procedure. Participants completed the same battery of measures as those administered in Study 1. The student interviewers underwent several training sessions designed to standardize procedures and most importantly, to ensure that the couples completed the measures independently of each other.

Results and General Discussion

The university women perceived they were in stronger relationships as a result of their partners' increased involvement in sports, although their closeness and affection for one another was independent of that involvement. Even so, the possibility remains that other aspects of their relationships were enhanced in ways not tapped by the RCI and *Affect for Partner Index*, e.g., trust, self-confidence.

Unlike the university sample, the older women in Study 2 were under no illusions that their partner's sporting interests were drawing them closer together. Table 1 shows that the extent of their male partners' involvement in sports was unrelated to their estimates of relationship strength, the RCI and *Affect for Partner* ratings.

Unlike women, older men saw their partners' involvement in sports as strongly strengthening their relationship. Additionally, the extent of female involvement was positively related to RCI scores. In the eyes of men, their partners' involvement in sporting activities is seen subjectively not only to strengthen their relationships but also by a standardized measure, to increase their closeness.

To summarize, with respect to the sport widow(er) hypothesis we see a perception of strengthened rather than weakened relationships among young females and older males. Additionally, the involvement in sports by the partners of older men is associated with closer relationships. In the main, relationships stand to be enriched by a partner's enthusiasm for sports.

REFERENCES

Bersheid, E., Snyder, M., & Omoto, A. M. (1989). The *Relationship Closeness Inventory*: Assessing the closeness of interpersonal relationships. *Journal of Personality and Social Psychology, 57,* 792-807.

Gantz, W., Wenner, L. A., Carrico, C., & Knorr, M. (1995a). Assessing the football widow hypothesis: A coorientation study of the role of televised sports on long-standing relationships. *Journal of Sport & Social Issues, 19,* 352-376.

Gantz, W., Wenner, L. A., Carrico, C., & Knorr, M. (1995b). Televised sports and marital relationships. *Sociology of Sport Journal, 12,* 306-323.

Quirk, K. (1997). *Not now honey, I'm watching the game.* New York: Fireside.

Roloff, M. E., & Solomon, D. H. (1989). Sex typing, sports interests, and relational harmony. In L. A. Wenner (Ed.), *Media, sports, and society* (pp. 290-311). Newbury Park, CA: Sage.

Smith, G. J., Patterson, B., Williams, T., & Hogg, J. (1981). A profile of the deeply committed male sports fan. *Arena Review, 5,* 26-44.

Wann, D. L., Melnick, M. J., Russell, G. W., & Pease, D. G. (2001). *Sport fans: The psychology and social impact of spectators.* New York: Routledge.

Effect of Different Normal/bizarre Image Combinations on Pair-cued Associated Word Recall

Alfredo Campos
University of Santiago de Compostela

María José Pérez-Fabello & María Calado
University of Vigo

The use of bizarre images as aide-mémoires is widely recommended in the psychology literature, although experimental studies have in fact provided little evidence that bizarre images are more effective than normal images. The present study investigated whether immediate and long-term associated word recall is affected by the subject's imaging ability or by the combination of images used. Specifically, subjects were presented with 10 pairs of words, in each case followed by an image-inducing sentence (normal-normal, normal-bizarre, bizarre-normal, or bizarre-bizarre), or with no image; each subject was presented with only one type of image-pairing sentence. Recall of the second item of each pair (cued by the first item) was tested both immediately and one week later. The subjects also completed a questionnaire to assess imaging ability. The results indicate that imaging ability had a significant effect on immediate recall, but not on one-week recall. The normal-normal image pairing protocol was more effective than the other protocols, both for immediate and one-week recall.

The use of bizarre mental images as aides-mémoire is an ancient practice: throughout history, writers on memory have advocated the use of unusual or bizarre images to help "fix" information (Campos & Pérez, 1996). This recommendation remains widespread, despite the fact that many experimental studies have not clearly confirmed that bizarre images are any more effective than normal images (see McDaniel & Pressley, 1987, for reviews).

Various factors appear to influence the effectiveness of bizarre images (Einstein & McDaniel, 1987; Wollen & Margres, 1987). First, study design: subjects may be presented with a list of words and be required to generate only normal images or only bizarre images (pure list design), or they may be required to generate both normal and bizarre images (mixed list design). Second, retention interval: recall may be tested immediately (short-term recall) or after a period of a week or more

---

Author info: Correspondence should be sent to: Alfredo Campos, Facultad de Psicología, University of Santiago de Compostela, 15782 Santiago de Compostela, Spain. E-mail: pscampos@usc.es

North American Journal of Psychology, 2002, Vol. 4, No. 2, 253-262.

© NAJP

(long-term recall). Third, recall type: subjects may be presented with a list of single words and be required to recall each word (free recall), or they may be presented with a list of associated pairs of words and be required to recall the second of each pair given the first (cued recall).

In most studies of immediate recall of associated pairs, bizarre images have proved ineffective. Specifically, bizarre images have proved ineffective for recall of associated word-pairs presented in pure lists (Bergfeld, Choate, & Kroll, 1982; Campos & Pérez, 1997; Iaccino, 1996; Riefer & LaMay, 1998) and for recall of associated word-pairs presented in mixed lists (Kroll, Jaeger, & Dornfest, 1992; Kroll, Schepeler, & Angin, 1986; Pra Baldi, De Beni, Cornoldi, & Cavedon, 1985; Riefer & Rouder, 1992). Riefer and Rouder (1992) found that bizarre phrases aided retrieval but not storage within memory. The positive effect of bizarre images on recall was reduced or eliminated when participants were presented with the first word of the associated pair, presumably because the cue word already acts to aid retrieval. However, bizarre images might aid *free* recall of paired word lists (i.e. recall without presentation of the first word of each pair).

In tests of immediate free recall, using pure lists, as opposed to immediate recall of associated pairs, the results are somewhat contradictory. Some authors have not found any significant differences between the effects of normal and bizarre images (Kline & Groninger, 1991; McDaniel & Einstein, 1986, 1991; Richman, Dunn, Kahl, Sadler, & Simmons, 1990; Robinson-Riegler & McDaniel, 1994). Other authors found better recall with normal images (Kline & Groninger, 1991; Kroll et al., 1992; Sharpe & Markham, 1992). Despite these conflicting results, most authors appear to agree that bizarre images are more effective than normal images for immediate free recall of words presented in mixed lists (Anderson & Buyer, 1994; Burns, 1996; Imai & Richman, 1991; McDaniel, Delosh, & Merritt, 2000; McDaniel, Einstein, DeLosh, May, & Brady, 1995; Mercer, 1996; Richman et al., 1990; Riefer & Rouder, 1992; Robinson-Riegler & McDaniel, 1994; Tess, Hutchinson, Treloar, & Jenkins, 1999).

The results of studies of long-term recall of associated pairs have likewise proved contradictory. Andreoff and Yarmey (1976) found that presentation of words with bizarre images improved one-day recall, using a mixed-list incidental-learning protocol. Other authors using mixed lists have not detected any difference between bizarre and normal images in recall after a day, a week or a month (Poon & Walsh-Sweeney, 1981) or after one week (Campos & Pérez, 1997; Kroll et al., 1986). In other studies, however, bizarre images have been found to have an effect on free recall of mixed lists 2 days after presentation (Kroll & Tu, 1988; Zoller, Workman, & Kroll, 1989) or one week after presentation

(Iaccino, Dvorak, & Coler, 1989; Sharpe & Markham, 1992). The positive effect of bizarre images on recall was reduced or it disappeared entirely when subjects were presented with the first word of the associated pair, presumably because the cue word already acts to aid retrieval. However, bizarre images might aid *free* recall of paired word lists (i.e. recall without presentation of the first word of each pair).

Mental images are nowadays widely considered to be one of the best ways of organizing information in memory, whether in temporary or permanent terms (Mayor & Moñivas, 1992). In general, direct instructions as to image formation, and any stimulus characteristic that facilitates image elicitation, have positive effects on recall (Bajo & Cañas, 1991).

Mental images play a particularly important role in mnemonic "systems," since associations are typically visual. Many studies that have analyzed the effectiveness of mental images in the recall process have found that "strong imagers" learn and recall more effectively than "weak imagers." Hiscock (1976), Denis (1982), Cornoldi and De Beni (1984), and Ernest (1991) have all found that strong imagers show better recall than weak imagers.

O'Brien and Wolford (1982) found significant differences between strong and weak imagers in the short-term recall of associated pairs, but not in long-term (one-week) recall. Campos and Pérez (1996) used a protocol in which half of the word-pairs were strong image-inducers, and the other half weak image-inducers; they found that strong imagers showed significantly better recall of short lists, but not of long lists. In a similar study, Campos and Pérez (1997) found that subjects' imaging ability had no significant effect on either short- or long-term recall of associated word-pairs, whether presented in short or long lists. However, Campos and Fernández (1997) found that subjects' imaging ability influenced serial recall of both short and long lists. In general, it seems that the capacity of subjects to form mental images influences immediate recall. The long-term effects are less clear.

In the present study we set out to determine whether subjects' imaging ability and the type of image-pairing protocol (bizarre-normal, normal-bizarre, bizarre-bizarre, normal-normal, or no-image) influence immediate or delayed recall of associated word-pairs presented in short lists. To the best of our knowledge, this is the first study to have compared all normal-bizarre pairings possible, using a protocol in which each subject received a list with only one type of pairing. In previous studies, subjects have received mixed lists, comprising word-pairs with normal image instructions and word-pairs with bizarre image instructions: in the present study, by contrast, subjects received lists, with all word pairs being presented either with normal-normal, normal-bizarre, bizarre-bizarre, or bizarre-normal image instructions, or with no-image instruction (i.e. rote method only). For example, the normal-normal image instruction for the pair *boy-cat* was "a boy chasing a cat", while the normal-bizarre image instruction for this pair was "a boy chasing a cat with wings".

METHOD

Participants

The sample comprised 267 participants, 138 male and 129 female, with a mean age of 16.2 years (range 14 - 19 years), selected at random from three public schools. All participants were urban-resident native speakers of Spanish who participated voluntarily in the study.

Instruments

All participants completed the *Vividness of Visual Imagery Questionnaire* (VVIQ) (Marks, 1973). This questionnaire comprises 16 items that are each scored twice, first with eyes open and then with eyes shut. These items measure the subject's ability to form images of people and objects (for example, "the colour and shape of trees"). Each item is evaluated on a five-point scale, with 1 indicating that the image is "perfectly clear and as vivid as normal vision," and 5 indicating "no image at all, you only know that you are thinking of the object." Note that high scores indicate weak imaging ability.

The 20 words were randomly selected from among those used in previous studies (Campos, 1989; Campos & González, 1992; Campos, Pérez-Fabello, & González, 2001), with the sole requirement that they should have high scores on imagery (score greater than 3.5 on a 1-7 scale). The mean imagery score of the 20 words selected was 6.53 (SD = 3.05); mean concreteness was 6.61 (SD = 2.31), and mean emotivity 3.46 (SD = 4.11). Subsequently, the words were paired at random to give a 10-pair list. We used a 10-pair list because 10 - 12 pairs is characteristic for short lists.

Procedure

First, the participants were randomly distributed among the five groups (bizarre-normal, normal-bizarre, bizarre-bizarre, normal-normal, and no image). Subsequently, in groups of 20 - 30, during normal class time, the participants were presented with the word list and asked to memorize the words by the strategy given (i.e. no images; or by imagery as induced by the sentence read after each word-pair, as described in what follows). The subjects in the no-image group (i.e. subjects who received a word list but no image instructions) were instructed to learn the words of the list by the rote method. The word list was presented aurally as a recording, at a rate of one word-pair every 15 seconds. After each word-pair, in all groups except the no-image control group, a sentence containing the two words was read, invoking normal and normal imagery, or normal and bizarre imagery, or bizarre and bizarre imagery, and bizarre and normal imagery, as appropriate. The sentences were of similar length, and each target word had at most one qualifier.

Once the word list had been presented, subjects received a sheet with the first word of each pair, and were given 90 seconds to write the second word of each pair. They then completed the VVIQ, with as much time as necessary. One week later, they performed the recall test again; subjects were not warned that they were going to repeat the task.

After the results had been processed, we classified subjects as strong or weak imagers, depending on whether their VVIQ score was below or above the mean score, respectively. Mean VVIQ score was 2.03 (*SD* = .53) for all subjects, 2.49 (*SD* = .42) for weak imagers, and 1.64 (*SD* = .24) for strong imagers.

RESULTS

To evaluate the influence of participants' image-forming ability and type of image instruction (normal-normal, normal-bizarre, bizarre-bizarre, bizarre-normal, no-image) on immediate cued recall, we used a two-factor analysis of variance (factors *imaging ability,* with 2 levels, and *instruction type*, with 5 levels). Mean immediate cued recalls in each group are shown in Table 1. As can be seen, imaging ability had a significant effect, $F_{(1,257)} = 6.112$, $p < .01$, as did image-pairing protocol, $F_{(4,257)} = 9.905$, $p < .001$. The two factors did not show significant interaction, $F_{(4,257)} = 1.499$, $p > .05$. Scheffé tests indicated significant

TABLE 1 Immediate Recall Scores (means and standard deviations) in each of the Five Groups, and in Participants Categorized on the Basis of Imaging Capacity (VVIQ score: note that high scorers are weak imagers, and low scorers strong imagers).

	M	*SD*	*N*
Normal-normal	9.04	.91	47
Normal-bizarre	7.08	2.42	49
Bizarre-normal	7.63	1.81	46
Bizarre-bizarre	7.62	1.76	69
No image	6.46	2.98	56
High VVIQ	7.04	2.46	113
Low VVIQ	7.89	2.02	154

pairwise differences only between the normal-normal group and the remaining groups.

A second two-factor ANOVA (factors *imaging ability*, with 2 levels, and *image instruction type*, with 5 levels) was performed to investigate effects of these variables on long-term cued recall. Mean long-term cued recalls in each group are shown in Table 2. Image-pairing protocol again had a significant effect, $F_{(4,257)} = 11.465$, $p < .001$, but neither imaging ability, $F_{(1,257)} = .911$, $p > .05$, nor the interaction between the two variables, $F_{(4,257)} = 1.868$, $p > .05$, was significant. Scheffé tests indicated a) that the normal-normal group achieved better recall than the normal-bizarre, bizarre-normal and no-image groups, and b) that the bizarre-bizarre group achieved better recall than the normal-bizarre group.

TABLE 2 Delayed Recall Scores (Means and Standard Deviations) in each of the Five Groups, and in Participants Categorized on the Basis of Imaging Capacity (VVIQ score: note that high scorers are weak imagers, and low scorers strong imagers).

	M	*SD*	*N*
Normal-normal	8.32	1.30	47
Normal-bizarre	5.22	2.62	49
Bizarre-normal	6.17	2.64	46
Bizarre-bizarre	7.06	2.62	69
No image	5.70	3.05	56
High VVIQ	6.20	2.88	113
Low VVIQ	6.73	2.62	154

DISCUSSION

For immediate recall of associated word-pairs, we found that normal-normal pairing was more effective than other types of pairing. In other words, bizarre images were less effective than normal images. This result coincides with that of most previous studies that have investigated the effect of bizarre images on immediate recall of associated word pairs, whether presented in pure lists (Bergfeld et al., 1982; Campos & Pérez, 1997; Iaccino, 1996; Riefer & LaMay, 1998) or mixed lists (Kroll et al., 1992; Kroll et al., 1986; McDaniel et al., 2000; Para Baldi et al., 1985; Riefer & Rouder, 1992; Tess et al., 1999).

For delayed recall of associated word-pairs, we likewise found that normal-normal pairing was more effective than bizarre-normal and normal-bizarre pairing, and than no-image. These results are again in agreement with those of previous studies of long-term recall (Campos & Pérez, 1997; Kroll et al., 1986; Poon & Walsh-Sweeney, 1981).

Our finding that subjects who used normal-normal images showed better recall than subjects who used other strategies can be interpreted in terms of the theory of "vertical" versus "horizontal" retrieval (Wollen & Margres, 1987). Vertical retrieval implies moving from one phrase to another, while horizontal retrieval implies retrieval of the key words of each phrase. Vertical retrieval depends on interconnection among sentences, whereas horizontal retrieval depends on interconnections within sentences. Cued recall consists primarily of horizontal retrieval. Bizarre images facilitate vertical retrieval, whereas normal images facilitate horizontal retrieval. Normal images show better integration (Wollen & Cox, 1981), and integration facilitates horizontal retrieval.

We found that participants' imaging ability influenced the immediate recall of associated word-pairs. This result coincides with that of previous studies, most of which have found that strong imagers show better recall than weak imagers (Campos & Fernández, 1997; Cornoldi & De Beni, 1984; Denis, 1982; Ernest, 1991; Hiscock, 1976). However, we did not detect any effect of imaging ability on long-term recall. O'Brien and Wolford (1982) likewise found that imaging ability affects short-term but not long-term recall.

Further studies are required to identify possible circumstances under which bizarre images are more effective aids to recall than normal images, and in what circumstances normal images are more effective than bizarre images. In addition to subject aptitudes like imaging ability, there is a clear need for more objective measures of imaging ability. It would be of interest to consider protocol variables like a) free recall versus associated pairs, b) pure versus mixed lists, c) immediate versus delayed recall, and d) imagery vividness of the words used.

REFERENCES

Anderson, D. C., & Buyer, L. S. (1994). Is imagery a functional component of the "bizarre imagery" phenomenon? *American Journal of Psychology, 107*, 207-222.

Andreoff, G. R., & Yarmey, A. D. (1976). Bizarre imagery and associative learning: A confirmation. *Perceptual and Motor Skills, 43*, 143-148.

Bajo, M. T., & Cañas, J. J. (1991). Las imágenes mentales. In J. J. Ruíz-Vargas (Ed.), *Psicología de la memoria* (pp.267-288). Madrid: Alianza.

Bergfeld, V. A., Choate, L. S., & Kroll, N. E. (1982). The effect of bizarre imagery on memory as a function of delay: Reconfirmation of the interaction effect. *Journal of Mental Imagery, 6*, 141-158.

Burns, D. J. (1996). The bizarre imagery effect and intention to learn. *Psychonomic Bulletin and Review, 3*, 254-257.

Campos, A. (1989). Emotional values of words: Relations with concreteness and vividness of imagery. *Perceptual and Motor Skills, 69*, 495-498.

Campos, A., & Fernández, C. I. (1997). Imagen mental e intervalo de retención en el recuerdo de series. *Estudios de Psicología, 58*, 105-112.

Campos, A., & González, M. A. (1992). Imagery, concreteness, emotionality, and meaningfulness values of words: Replication and extension. *Perceptual and Motor Skills, 74*, 691-696.

Campos, A., & Pérez, M. J. (1996). La estrategia de imágenes mentales extrañas como ayuda al recuerdo de pares asociados. *Apuntes de Psicología, 48*, 41-50.

Campos, A., & Pérez, M. J. (1997). Mnemonic images and associated pair recall. *Journal of Mental Imagery, 21*, 73-82.

Campos, A., Pérez-Fabello, M. J., & González, M. A. (2001). Propiedades de las palabras que influyen en su emotividad. *Revista de Psicología General y Aplicada, 54*, 179-191.

Cornoldi, C., & De Beni, R. (1984). Imagery and the "loci" mnemonic. *International Imagery Bulletin, 2,* 10-13.

Denis, M. (1982). Imaging while reading text: A study of individual differences. *Memory and Cognition, 10*, 540-545.

Einstein, G. O., & McDaniel, M. A. (1987). Distinctiveness and the mnemonic beneficits of bizarre imagery. In M. A. McDaniel & M. Pressley (Eds.), *Imagery and related mnemonic processes: Theories, individidual differences, and applications* (pp. 78-102). New York: Springer-Verlag.

Ernest, C. H. (1991). Ability differences and prose learning. *Intelligence, 15,* 455-477.

Hiscock, M. (1976). Effects of adjective imagery on recall from prose. *Journal of General Psychology, 94*, 295-299.

Iaccino, J. F. (1996). A further examination of the bizarre imagery mnemonic: Its effectiveness with mixed context and delayed testing. *Perceptual and Motor Skills, 83*, 881-882.

Iaccino, J. F., Dvorak, E., & Coler, M. (1989). Effects of bizarre imagery on the long-term retention of paired associates embedded within variable contexts. *Bulletin of the Psychonomic Society, 27*, 114-116.

Imai, S., & Richman, C. L. (1991). Is the bizarreness effect a spatial case of sentence reorganization. *Bulletin of the Psychonomic Society, 29*, 429-432.

Kline, S., & Groninger, L. D. (1991). The imagery bizarreness effect as a function of sentence complexity and presentation time. *Bulletin of the Psychonomic Society, 29*, 25-27.

Kroll, N. E. A., Jaeger, G., & Dornfest, R. (1992). Metamemory for the bizarre. *Journal of Mental Imagery, 16*, 173-190.

Kroll, N. E. A., Schepeler, E. M., & Angin, K. T. (1986). Bizarre imagery: The misremembered mnemonic. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12*, 42-53.

Kroll, N. E. A., & Tu, S. F. (1988). The bizarre mnemonic. *Psychological Research, 50*, 28-37.

Marks, D. F. (1973). Visual imagery differences in the recall of pictures. *British Journal of Psychology, 64*, 17-24.

Mayor, J., & Moñivas, A. (1992). Representación mental e imágenes mentales: II. Las imágenes mentales. In J. Mayor & M. De Vega (Eds.), *Memoria y representación* (pp.531-605). Madrid: Alhambra.

McDaniel, M. A., DeLosh, E. L., & Merritt, P. S. (2000). Order information and retrieval distinctiveness: Recall of common versus bizarre material. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 1045-1056.

McDaniel, M. A., & Einstein, G. O. (1986). Bizarre imagery as an effective memory aid: The importance of distinctiveness. *Journal of Experimental Pyshcology: Learning, Memory, and Cognition, 12*, 54-65.

McDaniel, M. A., & Einstein, G. O. (1991). Bizarre imagery: Mnemonic benefits and theoretical implications. In R. H. Logie & M. Denis (Eds.), *Mental images in human cognition* (pp. 183-192). Amsterdam: North Holland.

McDaniel, M. A., Einstein, G. O., DeLosh, E. L., May, C. P., & Brady, P. (1995). The bizarreness effect: It's not surprising, it's complex. *Journal of Experimental Psychology: Learning, Memory and Cognition, 21*, 422-435.

McDaniel, M. A., & Pressley, M. (1987). *Imagery and related mnemonic process*. New York: Springer.

Mercer, C. (1996). The bizarre imagery effect on memory. *Journal of Mental Imagery, 20*, 141-152.

O'Brien, E. J., & Wolford, C. L. R. (1982). Effect of delay in testing on retention of plausible versus bizarre mental images. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8*, 148-152.

Poon, L. W., & Walsh-Sweeney, L. (1981). Effects of bizarre and interacting imagery on learning and retrieval of the aged. *Experimental Aging Research, 7*, 65-70.

Pra Baldi, A., De Beni, R., Cornoldi, C., & Cavedon, A. (1985). Some conditions for the occurrence of the bizarreness effect in free recall. *British Journal of Psychology, 76*, 427-436.

Richman, C. L., Dunn, J., Kahl, G., Sadler, L., & Simmons, K. (1990). The bizarre sentence effect as a function of list length and complexity. *Bulletin of the Psychonomic Society, 28*, 185-187.

Riefer, D. M., & LaMay, M. L. (1998). Memory for common and bizarre stimuli: A storage-retrieval analysis. *Psychonomic Bulletin and Review, 5*, 312-317.

Riefer, D. M., & Rouder, J. N. (1992). A multinomial modeling analysis of the mnemonic benefits of bizarre imagery. *Memory and Cognition, 20*, 601-611.

Robinson-Riegler, B., & McDaniel, M. A. (1994). Further constraints on the bizarreness effect: Elaboration at encoding. *Memory and Cognition, 22*, 702-712.

Sharpe, L., & Markham, R. (1992). The effect of the distinctiveness of bizarre imagery on immediate and delayed recall. *Journal of Mental Imagery, 16*, 211-220.

Tess, D. E., Hutchinson, R. L., Treloar, J. H., & Jenkins, C. M. (1999). Bizarre imagery and distinctiveness: Implications for the classroom. *Journal of Mental Imagery, 23*, 153-170.

Wollen, K. A., & Margres, M. G. (1987). Bizarreness and the imagery multi-process model. In M. A. McDaniel & Pressley (Eds.), *Imagery and related mnemonic processes: Theories, individual differences, and applications* (pp.103-127). New York: Springer-Verlag.

Wollen, K. A., & Cox, S. D. (1981). Sentences cuing and the effectiveness of bizarre imagery. *Journal of Experimental Psychology: Human, Learning, and Memory*, 7, 386-392.

Zoller, C. L., Workman, J. S., & Kroll, N. E. A. (1989). The bizarre mnemonic: The effect of retention interval and mode of representation. *Bulletin of the Psychonomic Society, 27*, 215-218.

Explaining Punitiveness: Right-Wing Authoritarianism and Social Dominance

Jason S. Capps
University of Kansas

By administering a survey to 463 introductory sociology students, this paper introduces a Punitiveness Orientation Scale (PUN). Punitiveness—as a personality trait—is associated with punishment and defined as choosing harsh rather than lenient treatment of individuals. Additionally, the triangulation of right-wing authoritarianism (RWA), social dominance orientation (SDO), and punitiveness (PUN) are posited as being significantly interconnected. More importantly, punitiveness is shown to correlate even stronger with SDO (.49). Finally, when placed in a regression analysis it is shown that when coupled together RWA and SDO explain a significant amount of variance in punitiveness (31.9%).

Social scientists have often noted that over the last twenty years so-called *get tough* measures have dominated United States crime policy. Offenders are being given longer prison sentences as opposed to less punitive options, such as probation, fines, or community service. For instance, at the end of 1998, some 5.9 million people were either on probation, in jail, in prison, or on parole (U.S. Bureau of Justice Statistics, 2000). Also, state and federal prison authorities had 1,333,561 inmates under their jurisdiction at midyear 1999 (1,254,577 were physically in their custody) (U.S. BJS, 2000). Perhaps most tellingly, between 1990 and 1999 the incarcerated population grew an average of 5.8% annually, with the federal prison population rising 9.9% in one twelve month period alone (June 30 1998-June 30 1999) (U.S. BJS, 2000). Also, in the federal court, defendants are being sent to prison at an increasing rate. For example, in 1980 those convicted of a crime and sentenced to a federal prison numbered about 13,000; this compared to some 44,000 in 1998 (U.S. BJS, 1999).

Moreover, polling data routinely show that the public at large is in agreement with the idea that prison and the use of the death penalty are the best answers for curbing criminal behavior (Donziger, 1996; Rossi & Berk, 1997; Sherwood, Van Alstyne, & Dunn, 1988; Stinchcombe, Adams, Heimer, Scheppele, Smith & Taylor, 1980). A 1999 Gallup Poll, for instance, revealed that 71% were in favor of the death penalty for

---

Author info: Correspondence should be sent to: Jason S. Capps, Sociology Dept., University of Kansas, 716 Fraser Hall, Lawrence, KS 66045. E-mail jsc@ku.edu.

North American Journal of Psychology, 2002, Vol. 4, No. 2, 263-278.
© NAJP

murder. Also, it is well known that judges have tremendous latitude for imposing more lenient or harsh punishments in the sentencing phase of criminal trials. When faced with a sanctioning predicament, often judges are responsive to public opinion in deciding which sentences to inflict (Stalans & Diamond 1990; Walker & Marsh 1984). As the United States continues on this course of predominantly punitive measures it is essential to understand factors contributing to this particular ideology. This paper advocates the use of a new psychodynamic punitiveness scale for understanding more concretely the concept of punitiveness, prejudice, and other related personality traits. The consequences of punitive attitudes are vitally important in regard to public policy because these attitudes contribute to a mean-spirited and hostile public (Bray & Noble 1978; Durnham, Simpson, Rossi, & Miller, 1986; Frank, Cullen, & Cullen, 1987; Marcus, Sullivan, Theiss-Morse, & Wood, 1995; Rose & Prell, 1975). Moreover, the psychological propensity for punitiveness is often transferred to widespread support of overly punitive justice policies—three strikes laws, mandatory minimum sentencing, and increased use of the death penalty, to name a few. We need to understand what attitudes and beliefs contribute to this increase in the use of, and support for, punitive measures in an effort to advocate peaceful, humane, and sensible criminal justice policies.

Additionally, there is a gap in the literature on punitive attitudes due to the fact that appropriate assessment scales have not been constructed to directly study these tendencies. Further, there is a need to cultivate a punitiveness scale that will significantly aid, not only in our understanding of its relationship to authoritarianism, but also in our knowledge of more general prejudice attitudes. Another important reason for this research is to understand punitiveness as a personality trait that is not only related to criminal sanctioning, but also to attitudes toward the physical punishment of children and overall harshness. If a reliable PUN scale can be cultivated, then it is conceivable that significant progress is likely for theories of authoritarianism and prejudice. This study will substantiate many of Altemeyer's findings and encourage the use of a new scale to understand the complex nature of prejudice (Altemeyer 1996, 1998a, 1998b). The four main hypotheses tested in this essay are as follows: (1) A significant relationship exists between right-wing authoritarianism and punitiveness, (2) Significant differences just exist between women and men on the RWA, SDO, PUN, and PP-MAD, (3) RWA is presently the best psychodynamic predictor of punitiveness, (4) SDO is an independent and important variable in research on RWA and PUN.

Since 1970 Altemeyer has been involved in an ongoing project to understand and test Right-Wing Authoritarianism (RWA), which is close in definition to the original Adorno, Frenkel-Brunswick, Levinson and Sanford (1950) authoritarian personality type involving three crucial characteristics: conventionalism, submission, and aggression. Authoritarianism is a vitally important personality type and has been tested and shown as reliable in numerous studies (Adorno et al., 1950; Billings & Guastello, 1993; Christie, 1991; Kerlinger & Rokeach, 1966; Meloen, 1991,1994; Ray, 1983; Thompson & Michel, 1971; Zwillenberg, 1983).

In the latest studies employing the RWA Scale, (1994-1998) Altemeyer reports very strong alphas of .84 or above (Eigenberger, 1994; Leak & Randall, 1995; Ray, 1983; Rubinstein, 1996; Skitka & Tetlock, 1993; Stone & Smith, 1993). Unfortunately, like Adorno et al. (1950), Altemeyer has never developed a punitiveness scale. However, Christie (1993) reports a small part of Altemeyer's work that addresses the relationship between authoritarianism and punitiveness. As cited in Christie (1993), Altemeyer (1981) constructed vignettes consisting of five criminal scenarios and found positive linear correlations between RWA and criminal punitiveness. Basically, if individuals scored "high" on the RWA scale they were also likely to suggest the most severe punishment for offenders in the vignettes.

Moreover, Altemeyer has reported a .48 correlation between RWA and severity of sentencing (1993, p. 106). Over the years Altemeyer has also given his participants various scenarios involving individuals convicted of crimes and ascertained that high scoring RWAs are more likely to favor tougher punishment options. His work is, by far, the most extensive to date; however a few other researchers have also reported a positive correlation between these two constructs (Bray & Noble, 1978; Friend & Vinson, 1974; Giffitt & Garcia, 1979; Jurow, 1971; Rychman, Burns, & Robbins 1986).

I have constructed a Punitiveness orientation scale to further our understanding of punitive trends. Punitiveness, in this study, consists of a distinct personality trait relating to the treatment of children/physical punishment, advocating severe punishments for those convicted of crimes, and a general tendency to opt for punitive actions over leniency. The PUN scale includes statements like: (1) Spanking is often the most effective way to teach children not to hit others, (2) If your teenagers use drugs, you should turn them in to the police, and (3) If I were a juror I wouldn't hesitate to cast the decisive vote to send a murderer to death row. The definition of punitiveness in this paper is similar, as a construct, to what has been contrived in vignettes in past research (Altemeyer 1981, 1988, 1996, 1998a; Christie, 1993, Ercolani, Caprara, & Areni, 1977; Rychman, et al., 1986).

Altemeyer and other researchers have also stressed that the concept of authoritarianism includes interesting variations of the main features of

RWA. The most notable extension taps the *socially dominating* element of some authoritarian personalities and is vitally important. This distinction has been cultivated in two new scales that are relevant and worth exploration to detrmine if they are linked to judicial punishment attitudes and/or a punitiveness orientation. Most importantly the *Social Dominance Orientation Scale* (SDO; Pratto, Sidanius, Stallworth, & Malle, 1994) focuses specifically on, "a general attitudinal orientation toward inter-group relations, reflecting whether one generally prefers such relations to be equal, versus hierarchical" (Pratto & Sidanius, 1997, p. 742). In addition, the SDO Scale was constructed to investigate how individuals view justice, equality, and equal opportunity in society. As the authoritarianism literature suggests, it is logical to assume that individuals who scored high on the SDO Scale would also tend to score high on the *Punitiveness Orientation Scale*.

Additionally, Altemeyer clearly points out that the SDO and RWA Scales are tapping different parts of authoritarianism by stating, "...[the] SDO scale was built to tap into something different from both authoritarianism and interpersonal dominance" (1998, p. 53). The scales, used by Altemeyer (along with others), have therefore ostensibly found very weak correlations between the SDO and RWA, usually anywhere from .08 to .38 (1998, p. 55). Moreover, preliminary results also indicate that, although not highly correlated with RWA, the SDO Scale does positively correlate with prejudice against homosexuals (.42), against blacks (.52), against women (.49), against social policies favoring these groups, favoring punitiveness, and favoring support of capital punishment (1998, pp. 59-62). Likewise, Altemeyer argues that high SDOs do not particularly endorse submission to authorities; nor do they show attributes related to conventionalism, which are virtual constants in high RWAs (1998, p. 62).

Subsequently, Altemeyer cites research by McFarland and Adelson (1996) that found when combining the SDO and RWA scales as one set of attitudes, they could account for 50% of overall prejudice; a multiple r =.64 (1998, p. 55). Additionally, in Altemeyer's 1996 study he summarizes, "The two tests barely interconnected (.21 and .07), [and] they explained different segments of overall punitiveness...[N]one of the other personality tests mattered much once these two scales [SDO and RWA] had their say..." (1998, p. 55). Essentially, if these types of personality characteristics are included in an analysis of authoritarianism, Altemeyer contends that they can explain "uncommon levels of prejudice" (1998, p. 61).

Further, in Altemeyer's *Personal Power, Meanness, and Dominance Scale* (PP-MAD; 1995), SDO is taken a step further by measuring the level of cruelty in individuals who are willing to use power to get what they want no matter what the cost(s). This twenty-item scale is intended to describe the world view of the participants by asking such pro-trait items like: (1) it is a mistake to interfere with the 'law of the jungle,' some people were meant to dominate others, (2) would you be cold blooded and vengeful, if that is what it took to reach your goals?, and (3) I will do my best to destroy anyone who deliberately blocks my plans and goals (1998, p. 74). In populations tested thus far, Altemeyer has found consistent correlations between SDO and PP-MAD, usually anywhere from .43 to .61 (1998, p. 66). At the same time, the correlation between RWA and PP-MAD has been very low, -.08 in one study (1998, p. 66).

METHOD

Participants

The sample consisted of 463 students in two introductory sociology courses. Eighty-five percent of the students surveyed were in their first or second year of college. Four hundred sixty three, i.e., 94% of the original sample of 495, indicated their sex on the survey (with 42% being male and 58% female). In addition, most students were between the ages of 18-23. The racial/ethnic distribution of my sample was: 78.3% White, 18.4% Hispanic, 5.0% Asian (Asian-American), 2.9% African-American, and 2.0% American Indian or Alaskan Native (with 48.3% disclosing).

Procedure

A valiant effort was made to find a solid measure of punitiveness using a Likert-type format. Zwillenberg's (1983) scale provided a sketch of important attributes of the construct, but lacked specificity. Given the success of the RWA and other similar scales it was determined that a punitiveness orientation scale might prove to be a productive measure. The survey booklet began with Altemeyer's (1996) Right-Wing Authoritarianism scale (RWA) followed by the Smith and Capps (1999) Punitiveness Orientation scale (PUN) shown in Appendix A. Also included were Pratto et al.'s (1994) Social Dominance Orientation scale (SDO) and Altemeyer's (1996) Personal Power, Meanness, and Dominance scale (PP-MAD). All scales were organized in a Likert-type format with equal numbers of pro-trait and con-trait items. The booklet ended by asking the respondents to specify their gender and age.

RESULTS

Table 1 specifies the properties of all scales used in this study. It is worth noting that the mean RWA score of 131 is remarkably similar to what Altemeyer has reported over the past several years (his mean routinely hovers around the 130-135 mark) (Altemeyer, 1996; 1998a).

TABLE 1 Psychometric Properties of Scales

	Mean	*SD*	*Alpha*	Mean Inter-Item *r*
PUN	99.1	12.9	.70	.13
RWA	130.7	41.4	.94	.32
SDO	63.1	22.5	.89	.32
PP-MAD	80.3	22.0	.82	.20

Note. PUN=*Punitive Orientation Scale* from 15-135; RWA=*Right-Wing Authoritarianism* from 33-297; SDO=*Social Dominance Orientation* from 14-126; PP-MAD=*Personal Power, Meanness, and Dominance* from 20-180.

As Table 1 shows, the RWA scale had an alpha of .94. Further, the factor analysis using the rotated varimax format found that the scale had two eigenvalues, namely a 10.92 and a 2.54 accounting for 42.1% of the cumulative variance. Interestingly, Altemeyer reported in a recent article (1999) eigenvalues of 9.53 and 2.42 for his most current 30-item RWA scale. Similarly, Altemeyer found that these two strong eigenvalues accounted for 32% and 40% of the test variance, respectively—almost identical to what is reported here (1999, pp. 5-6).

The SDO scale, meanwhile, had an alpha of .89 and had virtually one underlying dimension. Here, the eigenvalues were 6.19 and 1.60 respectively, accounting for 45.6% of the cumulative variance. The PP-MAD Scale had an alpha of .82 with eigenvalues of 6.80 and 2.36, covering 45.8% of the overall variance. Finally, the PUN scale had an alpha of .70. There were five eigenvalues over 1.0 when factor analysis was performed. However, since the top three factors accounted for 41% of the variance (2.94, 1.72, and 1.43 respectively), it can be reasonably assumed that the PUN scale is a fair first approximation of a general punitiveness orientation.

The most important conclusions regarding the quantitative data analysis are the following: (1) Validation of the traditional association between RWA and PUN in previous research, (2) Evidence suggesting significant differences in scaled scores between men and women on the PUN, SDO, and PP-MAD, (3) An indication that SDO is a relevant and independent source of PUN and (4) Speculation that perhaps SDO is even more important in explaining variance on PUN than RWA.

Before presenting the regression results, employing the dependent variable PUN, the correlation matrix of the four scales is presented. It is important to note the correlation between PUN with RWA and SDO. Moreover, implicit in the theoretical model is the admission that the independent variables are indeed to an extent inter-correlated with each other. It is obvious that the scales need to be administered to many more populations (especially the PUN scale) in order to find out if multicollinearity is a persistent problem that is worthy of further investigation. Nonetheless, the results of the correlations are listed below in Table 2.

TABLE 2 Correlations of Variables in the Model

	Measure			
	SEX	PP-MAD	SDO	RWA
PUN	-.17	.32	.49	.45
RWA	-.07	.14	.36	
SDO	-.19	.53		
PP-MAD	-.25			

Note. All correlations are significant at the .01 level.

The positive linear correlation between the PUN scale and RWA was .45. If individuals score high on the PUN scale, they are likely to score high on the RWA Scale. This finding is consistent with other research on these two variables in the last 30 years (Altemeyer, 1981, 1988, 1996, 1998a; Bray & Noble, 1978; Centers, Shomer, & Rodrigues, 1970; Friend & Vinson, 1974; Griffit & Garcia, 1979; Jurow, 1971; Zwillenberg, 1983).

Perhaps even more telling is the .49 correlation between PUN and SDO. Given that recent research has shown that there are at least two different forms of authoritarianism (RWA and SDO), it is imperative to understand if one construct is a more adequate predictor of PUN. By a narrow margin (.45—RWA and .49—SDO) the data suggest that SDO might be a stronger survey measure.

To get a better estimation, specification, and overall prediction of the relationship, a multiple regression analysis was done. However, before collating the regression tables, it should be pointed out that there were significant differences in scale scores depending on gender/biological sex—which I did not initially hypothesize— and therefore this topic will be investigated next.

It is striking to note that the mean differences in scores based on gender/biological sex were significant. Most researchers have argued that there are no significant sex differences in authoritarianism (Adorno, 1950; Altemeyer, 1981, 1988, 1998a, 1998b; Billings & Guastello, 1993; Peterson, Doty, & Winter, 1993; Thompson & Michel, 1971). In recent work with the RWA scale, for instance, Altemeyer and others report no significant sex differences. In regard to the SDO and PP-MAD scales, on the other hand, Altemeyer reports that high SDOs and high PP-MADs are more likely to be men. Also, since the PUN scale is being tested for the first time it is important to find out if significant differences exist.

All past studies employing the SDO Scale indicate that men are significantly higher scorers than women. My findings echo earlier studies that men tend to be more socially dominant. As Table 4 shows, sex was negatively correlated with the five main independent variables in every instance: PUN (-.17), RWA (-.07), SDO (-.19) and PP-MAD (-.25). In order to get a more complete picture of the potential differences in scores on the independent variables, *t*-tests were administered comparing the difference between the means of women and men. Table 3 describes the mean score differences on the PUN, RWA, SDO, and PP-MAD. As expected, there was not a significant difference in scores on the RWA scale. However, there were significant differences on the PUN, SDO, and PP-MAD, all of which were statistically significant at the .001 level. Nevertheless, it is safe to say that, in general, females were likely to be less punitive, less socially dominating, and less prone to personal power, meanness, and domination.

TABLE 3 Mean Score Differences on Variable by Sex

	Mean	SD	t	p-value
PUN Scale				
Male	101.7	13.7	3.65	.000***
Female	97.3	12.0		
RWA Scale				
Male	133.8	44.5	1.40	.164
Female	128.4	39.0		
SDO Scale				
Male	68.04	24.0	4.20	.000***
Female	59.41	21.0		
PP-MAD Scale				
Male	86.70	21.0	5.52	.000***
Female	75.60	22.0		

Note. The total number of respondents is 463, meaning 32 students did not reveal their sex on the survey, 1=Male 2=Female.
***$p<.001$; **$p<.01$; *$p<.05$.

Table 4 is a summary of the regression of PUN with the selected four independent variables. In this case when RWA was regressed on punitiveness orientation the unstandardized regression coefficient was .138, and it was statistically significant at the .001 level. If these two variables (RWA and PUN) were placed in the same linear regression equation, for every 10-point increase on the RWA scale there was a 1.38 point increase on the PUN scale. Also, when RWA was the only variable in the model it accounted for 19.6% of the overall variance in the general punitiveness orientation. It should also be pointed out that when SDO was added to RWA, the regression coefficient went to .216 (.000) and explained 31.9% of the variance in PUN. Obviously, this was a significant increase (an adjusted R^2 change of .123) and was subsequently the *best- fit* model in Table 4. In other words, when trying to explain variance in the general punitiveness orientation, RWA and SDO were the two best of the four independent variables included in my analysis. When PP-MAD and SEX were entered they added little to the overall explained variance. Also, it is worth noting that as independent variables were added (beyond RWA) the regression coefficient for RWA went down (.097, .099, .096, and .096) but was nevertheless still statistically significant.

TABLE 4 Summary of Punitiveness Regression Analysis for RWA, SDO, PP-MAD, and SEX (N=463)

	Variable	b	SE B
Step 1	RWA	.138***	.013
Step 2	RWA	.097***	.013
	SDO	.216***	.024
Step 3	RWA	.099***	.013
	SDO	.180***	.027
	PP-MAD	.065*	.026
Step 4	RWA	.096***	.013
	SDO	.170***	.028
	PP-MAD	.027	.031
	SEX	-1.74	1.03

Note. Adjusted R^2=.20 for Step 1; .32 for Step 2; .33 for Step 3; and .33 for Step 4. ***p<.001; **p<.01; *p<.05; b=unstandardized coefficient.

DISCUSSION

There are several main findings to emerge from this study. Given the thoroughness of Altemeyer's work on RWA, SDO, and PP-MAD over the past several years, it is imperative that my discussion begins by pointing out corroborations and discrepancies between my study and past work.

In regard to the correlation of RWA and PUN, Altemeyer reports on studies linking punitiveness with authoritarianism: "RWA scores have often predicted the length of prison terms hypothetically imposed upon persons convicted of crimes. Among Manitoba students the correlation for overall sentences in a ten-trial survey usually ran between .40 and .50" (1996, p. 22). Altemeyer concludes, "believing in punishment as they do [high RWAs] and enjoying dishing it out as they do, they are more likely in general to sentence wrongdoers to long prison terms" (1996, p. 23). My findings corroborate Altemyer's, given the positive correlation found between RWA and PUN.

Likewise, in comparing the RWA Scale with the SDO scale, Altemeyer reports a meager .08-.21 correlation in different populations (in an attempt to say that they are in fact measuring two different attitudinal tendencies—combating multicollinearity) (Altemeyer, 1996, p. 62). My results seem to suggest a bit otherwise, as Table 3 showed that there are, indeed, a couple of the scales that correlated higher than Altemeyer might predict. When specifically looking at the relationship between RWA and SDO, they are correlated at .36 in this data set. This finding diverges somewhat with Altemeyer's implication that these two measures are *totally* distinctive. However, it seems likely that there are clear differences between these groups. He lists a few attributes of an ordinary authoritarian (or a person with a high RWA Score) as being the following: fearful, self-righteous, fundamentalist, conforming to laws, justifying punishment, and tending to view themselves as kind and considerate. On the contrary a social dominator usually has the following distinctive characteristics: self-confident, a-moral, anti-religion, scornful of conformity, enjoying punishing others, and self-consciously rather cruel. My results verify Altemeyer's conjecture that SDOs and RWAs are indeed qualitatively distinctive. High SDOs are more prejudiced than high RWAs. Further, high SDOs seem to not *act* as aggressively in most cases as high RWAs, and high SDOs seem opposed to one of democracy's central values, overall *equality* (1998, p.82).

Moreover, Altemeyer (1998) argues that the PP-MAD scale and the SDO scale are usually moderately correlated with one another, when compared with the PP-MAD and SDO scales, which are *weakly* correlated with RWA. He asserts, "[In] my probing attempts to discover 'what makes SDOs run?' [he introduced] the entirely *ad hoc Personal Power, Meanness, and Dominance* scale. Summed PP-MAD scores correlated an unambiguous .61 with SDO, -.08 with RWA, and .42 with gender" (1998, pp. 74-75). Altemeyer is again making the point that there are distinct differences between high RWAs and high SDOs, and if an individual scores high on PP-MAD it may indicate whether or not he/she is a high RWA or high SDO. In addition, Altemeyer notes that in a parent study, "PP-MAD relationship with SDO obtained with students reappeared among these parents (.59). But, again, RWA scores were marginally correlated with being power-hungry, mean, and domineering (.15)" (1998, p. 76). In my sample the PP-MAD and SDO scales correlated at .53, which is similar to Altemeyer's findings. The link between PP-MAD and RWA was a low .14, again almost identical to what has been discovered in the past (Altemeyer, 1996; 1998a; 1998b).

As for the significant gender differences discovered for SDO, PP-MAD, and PUN in my data set; Altemeyer's (1998) data have also shown that men tended to be more socially dominant when compared to women. Further, Altemeyer has often noted that men tend to score higher on the PP-MAD as well. Perhaps the most interesting finding, in regard to gender differences, has to do with the *t*-score of the PUN measure. At this time it would be hard to even speculate why this is the case, but there is certainly the possibility for more theoretical research investigating what might lie behind this finding. In other words, why do men tend to be more punitive than women?

Furthermore, my results suggest that the RWA and SDO scales are jointly productive for explaining punitiveness, which is congruent with McFarland and Adelson's (1996) research in explaining another kind of prejudice: *ethnocentric* attitudes. Similarly, McFarland and Adelson also found that the two best predictors for explaining ethnocentrism were SDO and RWA, with SDO being perhaps the more important. Altemeyer also reports, "A regression analysis of the sum of Sam McFarland's four tests found that SDO and RWA *as a package* accounted for 50% of overall prejudice" (1998, p. 60). Again, my results indicate that SDO and RWA account for 32% of overall punitiveness. Altemeyer expounds on the importance of research involving SDO by saying, "The implications of McFarland and Adelson's discovery, should it reappear in other populations and with other measures, overpower one. For if you want to explain the many kinds of prejudice exposed in this situation, they are largely matters of personality. And only two kinds of personality are basically involved: the social dominator and the right-wing authoritarian" (1998, p. 60). My results echo the conclusion that RWA and SDO are the two main predictors of a different type of prejudice: punitiveness.

Additionally, the conclusion has been drawn that perhaps SDO is a more accurate predictor of PUN when compared with RWA or PP-MAD. This finding diverges somewhat from recent research by Altemeyer (1998). In one of the many pilot groups used in Altemeyer's work (including the RWA, SDO, and PP-MAD) he focused on the responses of 400 parents (of his psychology students at Manitoba). Imbedded in the various scales was a short set of items related to the proper criminal sanctions that should be imposed. In one of his groups parents were asked to impose a sentence for: spitting on a premier, being convicted of a string of muggings, and of dealing heroin (no mention of the status of the offender was given). Given these scenarios Altemeyer noted only a

.12 correlation with sentences handed down in the trial situation with the SDO Scale, as well as a stronger .32 correlation with the RWA Scale (1998, p. 66). Clearly, the RWA Scale seemed to be a better predictor of punitiveness in this particular sample. It is noteworthy that in my sample SDO was only slightly higher correlated with PUN (a .46—RWA and .49—SDO).

In another group parents were asked to evaluate the proper sentencing for, "either a panhandler or an accountant who attacked the other after a sidewalk argument" (1998, p. 67). Interestingly, in this case there was a significant difference in how the high RWAs and high SDOs responded, with high RWAs sentencing the panhandler to longer prison terms when compared with the accountant found guilty of the same crime. The high SDOs did not tend to favor the discrepancy in the sentences meted out to the two different offenders (1998, pp. 68-69). This pattern, in which high RWAs often have double standards, while SDOs do not, has been shown as a fairly significant pattern in many relevant studies (Altemeyer, 1996; 1998a; 1998b; Pratto, Sidanius, Stallworth, & Malle, 1994; Sidanius & Pratto, 1999). It is quite evident that more studies need to be done in order to make definitive statements regarding the tenuous relationships among RWA, SDO, and PUN.

It has been fifty-two years since Adorno and his associates published *The Authoritarian Personality*. Over that span there has been a significant body of research devoted to learning more about authoritarianism and the different nuances involved in the authoritarian personality. Research in this area suggests that these distinctive personality traits are patterned, complex, and a significant part of some people's personalities, therefore requiring scholars to persistently study anti-democratic tendencies—in the hope of understanding and ultimately alleviating these unwanted attitudes and beliefs. Further, it has been well documented that punitive and harsh treatment, in response to individuals who break the law, is the preferred way to handle these situations by the public. Given the widespread support of punitive measures it is vital to understand why these beliefs are so prevalent. I have posited the possibility that the tendency toward punitiveness is related to one's level of authoritarianism. In my sample I established a positive relationship between RWA, SDO, and PUN. Social psychologists also need to learn what social structural factors are behind these tendencies. For example, what roles do the family, peer relationships, religious beliefs, or a plethora of other social factors play in encouraging or discouraging authoritarianism, social dominance, and punitiveness? Also, the fact that significant gender differences existed in the SDO, PP-MAD, and PUN is worth further study to understand possible reasons why this is the case.

An even more pointed question is how to turn back the tide of punitiveness toward a more informed, peaceful, and socially just society. As previously noted, the criminal justice system is expanding wildly and the majority of the public seems to be on board. One way to fend off this knee jerk reaction to crime, of increased punitiveness, is to understand a punitive orientation and its correlates. Hopefully, the research presented has aided in our understanding of authoritarianism and punitiveness and posed interesting research questions for continued work in this area.

REFERENCES

Adorno, T.W., Frenkel-Brunswick, E., Levinson, D., & Sanford, R.N. (1982). Abridged Edition. *The authoritarian personality*. New York: W. W. Norton.

Altemeyer, B. (1981). *Right-wing authoritarianism*. Winnipeg: University of Manitoba Press.

Altemeyer, B. (1988). *Enemies of freedom: Understanding right-wing authoritarianism*. San Francisco: Jossey-Bass.

Altemeyer, B. (1996). *The authoritarian spector*. Cambridge, Massachusetts: Harvard University Press.

Altemeyer, B. (1998a). The other 'authoritarian personality.' In M.P. Zanna (Ed.), *Advances in Experimental Social Psychology* (Vol. 30, pp.48-92). New York: Academic.

Altemeyer. B. (1998b). The 20-item RWA scale. Unpublished paper: University of Manitoba: Winnipeg, Manitoba.

Billings, S.W., & Guastello, S.J. (1993). A comparative assessment of the construct validity of three authoritarian measures. *Journal of Research In Personality, 27*, 1424-1430.

Bray, R.M., & Noble, A.M. (1978). Authoritarianism and decisions of mock juries: evidence of jury bias and group polarization. *Journal of Personality and Social Psychology, 36*, 1424-1430.

Christie, R. (1993). Some experimental approaches to authoritarianism: authoritarianism and punitiveness. In W. F. Stone, G. Lederer, & R. Christie (Eds.), *Strength and weakness: The authoritarian personality today* (pp. 99-118). New York: Springer-Verlag.

Christie, R. (1991). California F scales. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 527-554). San Diego: Academic Press.

Donziger, S.R. (1996). The real war on crime: The report of the National Criminal Justice Commission. New York: Harper-Perennial.

Durhman, A.M., Simpson, J.E., Rossi, P. H., & Miller, J.L. (1986). The use of factorial survey design in assessments of public judgments of appropriate punishment for crime. *Journal of Quantitative Criminology, 2*, 181-190.

Eigenberger, M. (1994). Physique. temperament and beliefs: A philosophical and empirical exploration of W.H. Sheldon's constitutional psychology. Unpublished master's thesis, University of Wyoming, Laramie.

Ercolani, A.P., Caprara, G., & Areni, A. (1977). Effects of authoritarianism on attitudes toward crimes and criminals. *Giornale Italiano Di Psicologia, 4*, 313-320. Frank, J., Cullen, F.T., & Cullen, J.B. (1987). Sources of judicial attitudes toward criminal sanctioning. *American Journal of Criminal Justice, 2*, 151-164.

Friend, R.M., & Vinson, M. (1974). Leaning over backwards: Jurors' responses to defendants' attractiveness. *Journal of Communication 24*, 124-129.

Griffit, W., & Garcia, L. (1979). Reversing authoritarian punitiveness: the impact of verbal conditioning. *Social Psychology Quarterly, 42*, 425-427.

Jurow, G.L. (1971). New data on the effect of "death qualified" jury on the guilt determination process. *Harvard Law Review 84*, 567-611.

Kerlinger, F. & Rokeach, M. (1966). The factorial nature of the F and D scales. *Journal of Personality and Social Psychology, 4*, 391-399.

Leak, G.K. & Randall, B.A. (1995). Clarification of link between right-wing authoritarianism and religiousness: The role of religious maturity. *Journal for the Scientific Study of Religion 34*, 245-252.

Marcus, G.E., Sullivan, J.L., Theiss-Morse, E., & Wood, S.L. (1995). *With malice toward some: How people make civil liberties judgments.* Cambridge, Massachusetts: Cambridge University Press.

McFarland, S.G., & Adelson, S. (1996, July). An omnibus study of personality, values, and prejudice. Paper presented at the 19th meeting of the International Society of Political Psychology, Vancouver, British Columbia.

Meloen, J.D. (1994). A critical analysis of forty years of authoritarianism research: did theory testing suffer from cold war attitudes? In Russell F. Farnen (Ed.), *Nationalism, ethnicity, and identity: cross national and comparative perspectives* (pp. 127-165). New Brunswick: Transaction.

Meloen, J.D. (1991). The fortieth anniversary of the authoritarian personality. *Politics and the Individual, 1*, 119-127.

Peterson, B.E., Doty, R.M., & Winter, D.G. (1993). Authoritarianism and attitudes toward contemporary social issues. *Personality and Social Psychology Bulletin, 19*, 174-184.

Pratto, F., Sidanius, J., Stallworth, L.M., & Malle, B.F. (1994). Social dominance orientation: A personality variable predicting social and political attitudes. *Journal of Personality and Social Psychology, 67*, 741-763.

Ray, J. J. (1983). Defective validity in the Altemeyer authoritarianism scale. *The Journal of Social Psychology, 125*, 271-272.

Rose, A. M., & Prell, A. E. (1975). Does the punishment fit the crime? A study in social valuation. *American Journal of Sociology, 1*, 247-259.

Rossi, P.H., & Berk, R.A. (1997). *Just punishments—federal guidelines and public views compared.* Hawthorne, New York: Walter de Gruyter.

Rubinstein, G. (1996). Two peoples in one land: A validation study of Altemeyer's right-wing authoritarianism scale in the Palestinian and Jewish societies in Israel. *Journal of Cross-Cultural Psychology 27*, 216-230.

Rychman, R. M., Burns, M. J., & Robbins, M. A. (1986). Authoritarianism and sentencing strategies for low and high severity crimes. *Personality and Social Psychology Bulletin, 12*, 227-235.

Sherwood, Z.E., Van Alstyne, D.J., & Dunn, C.S. (1988). The national punishment survey and public policy consequences. *Journal of Research in Crime and Delinquency, 25*, 120-149.

Sidanius, J., & Pratto, F. (1999). *Social dominance: an intergroup theory of social hierarchy and oppression*. UK, Cambridge: Cambridge University Press.

Skitka, L.J. & Tetlock, P.E. (1993). Providing public assistance: Cognitive and motivational processes underlying liberal and conservative policy preferences. *Journal of Personality and Social Psychology, 65*, 1205-1223.

Stalans, L.J., & Diamond, S.S. (1990). Formation and change in lay evaluations of criminal sentencing: misperception and discontent. *Law and Human Behavior*, *14*, 199-214.

Stinchcombe, A.L., Adams, R., Heimer, C.A., Scheppele, K.L., Smith, L., & Taylor, D.G. (1980). *Crime and punishment—changing attitudes in america*. San Francisco, California: Jossey-Bass.

Stone, W.F., & Smith, L.D. (1993). Authoritarianism: Left and right. In Stone, W.F., Lederer, G. & Christie, R. (Eds.), *Strength and weakness: The authoritarian personality today*. New York: Springer-Verlag.

Thompson, R.C., & Michel, J.B. (1971). Measuring authoritarianism: A comparison of the F and D Scales. *Journal of Personality and Social Psychology, 1*, 180-190.

U.S. Department of Justice, Bureau of Justice Statistics. (1998). Bureau of justice statistics: homicide trends. www.ojp.usdoj.gov/bjs/.

U.S. Department of Justice (1999). Bureau of justice statistics: trends in federal cases. www.ojp.usdoj.gov/bjs/.

U.S. Department of Justice (2000). Bureau of justice statistics: corrections statistics.www.ojp.usdoj.gov/bjs/.

Walker, N. & Marsh, C. (1984). Do sentences affect public disapproval? *British Journal of Criminology, 24*, 27-48.

Zimmerman, S., Van Alstyne, D.J., & Dunn, C.S. (1988). The national punishment survey and public policy consequences. *Journal of Research in Crime and Delinquency, 25*, 120-149.

Zwillenberg, D.F. (1983). *Predicting biases in the punishment of criminals as a function of authoritarianism: the effects of severity of the crime, degree of mitigating circumstances, and status of the offender*. Ph.D. Dissertation, Columbia University.

APPENDIX A:
2000 PUNITIVENESS ORIENTATION SCALE (SMITH AND CAPPS)

1. It's unreasonable to give people stiff prison sentences simply for possessing small quantities of drugs for personal use.*
2. In most cases probation is simply an unjustified way of putting criminals back on the street.
3. The death penalty is never an appropriate punishment, even for murder.*
4. Three-time losers deserve to be sentenced to life without the possibility of parole.
5. Spanking is often the most effective way to teach children not to hit others.
6. Punishment simply for the purpose of getting revenge is unacceptable.*
7. The courts should do everything they can to prevent law enforcement officers from physically harming or intimidating crime suspects.*
8. Physically punishing misbehaving children may hurt them in the short run, but it will help them in the long run.
9. Teachers should be forbidden to physically punish children who misbehave.*
10. I would never personally throw the switch to execute a condemned prisoner, no matter what his crime might have been.*
11. I think private citizens should take matters into their own hands if the courts are unwilling to punish criminals properly.
12. People should never kick or hit their pets.*
13. If children refuse to eat what their parents serve them, they should be required to stay at the table until they change their minds.
14. If your teenagers use drugs, you should turn them in to the police.
15. If I were a juror I wouldn't hesitate to cast the decisive vote to send a murderer to death row.

*Con-trait item, for which the –4 to +4 scoring key is reversed.

Authors Note: This paper was presented in Seattle, WA at the Annual Meeting of the International Society of Political Psychology (July, 2000). I would especially like to thank Sam McFarland for his critique of the manuscript and David N. Smith for aiding in the process of constructing our punitiveness orientation scale.

Status, Apparel and Touch: Their Joint Effects on Compliance to a Request

Nicolas Guéguen
Université de Bretagne-Sud Vannes - France

Research shows that touch can increase compliance to a request. In this study we investigated how a higher social status, operationalized by the apparel of the solicitor, can also enhance compliance. Participants in an experiment were solicited for a street survey. The interviewer (a man) was clothed in a style conveying a person of high, low or intermediate status. While formulating his request, he touched (or did not touch) the respondent. Results show that higher status associated with touch increased compliance to the request.

Touching the arm or shoulder of a person for 1 to 2 seconds when asking for a favor seems trivial. However, this brief nonverbal contact significantly influences compliance. Brockner, Pressman, Cabitt and Moran (1982) showed that a request to return a dime left in a phone booth accompanied by a light touch on the arm increased compliance from 63 % (no-touch control situation) to 96 %. Similarly, when asked for a dime, passers-by complied in 51 % of the cases when touch was used and 29 % of the cases when no-contact was made (Kleinke, 1977a). In the same way, the percentage of a petition signing increased from 55 % (no touch control condition) to 81 % when a slight touch was made during the request (Willis & Hamm, 1980). Hornik and Ellis (1988) found that touch increased compliance for an interview. Hornik (1987) observed that touch increased the number of persons responding to a street survey and increased compliance in answering a subsequent questionnaire.

Also, when touched, people are more persistent when executing a difficult task consisting of answering a long questionnaire on very provocative subjects (Nannberg & Hansen, 1994). Likewise, a simple touch of a client by a waiter or a waitress in a restaurant increases the amount of the tip (Crusco & Wetzel, 1984; Hornik, 1992b; Lynn, Le & Sherwyn, 1998; Stephen & Zweigenhaft, 1986). Willingness to taste products increases when an employee touches shoppers when making the request and increases the selling rate of the product (Hornik, 1992a; Smith, Gier & Willis, 1982). In the same way, Kaufman and Mahoney

---

Author info: Correspondence should be sent to: Dr. Nicolas Gueguen, Universite de Bretagne-Sud Vannes, France.

North American Journal of Psychology, 2002, Vol. 4, No. 2, 279-286.

© NAJP

(1999) showed that when touched by a waitress, patrons of a public tavern consumed more alcohol than patrons who were not touched.

Helping behavior is also influenced by touch. Goldman and Fordyce (1983) found that when people were touched by a confederate during an interview, greater helping behavior was observed toward a confederate who dropped several questionnaires on the ground. When touched, people were more likely to answer telephones for a charity telethon for crippled children (Goldman, Kiyohara & Pfannensteil, 1985).

Given this research showing that touch has a positive effect on compliance to a request, the question arises as to why this effect occurs in interactions among strangers. Research relating to the initiation of touch between people could be used to explain this influence. Research shows that touch is often initiated by individuals of higher status towards individuals of lower status (Hall, 1996; Henley, 1973). The role of touch as a status indicator has support in the literature (Major & Heslin, 1982; Summerhayes & Suchner, 1978). These experimenters asked their subjects to examine a series of still photographs portraying dyadic interactions. In half of the cases, one person is obviously touching the other. The results show that in the touch condition, the "toucher" was evaluated as more dominant than the "touchee," whereas no perceived difference in dominance was found between the two persons when no contact occurred. From these results we can speculate that when a stranger initiates touch while asking for help that stranger is perceived as having a higher status. Although this assumption has not been tested directly with touch, related research has shown that a solicitor wearing high status clothing is likely to experience compliance with his or her request (Goodman & Gareis, 1993).

Numerous studies show that high status people are likely to get help when they solicit it. In most of these studies, the status was manipulated either by the apparel of the solicitor or by staus-relevant information given to participants. Kleinke (1977b) showed that people will more readily give a coin to a stranger in the street if the individual is neatly dressed (80 %) rather than to someone whose appearance is slovenly (32 %). These results were more recently confirmed by an experiment conducted by McElroy and Morrow (1994) when the solicitation was aimed at raising money for a charitable organization. Status is also linked with the individual's integrity. Thus, Bickman (1971) has demonstrated that participants would more readily give back the money they had found in a telephone booth to a high status individual when this person argued that it was his/hers, rather than when it was an individual of equal status. These studies highlight an effect of status upon compliance to requests emanating from high status individuals. Nevertheless, the same effect was observed on spontaneous helping behavior. Solomon and Herman

(1977) found that passersby were more willing to help a high-status woman than a mid-status woman who had dropped her groceries while loading them in the trunk of her car.

To my knowledge, no research has tested the joint effect of status and touch on compliance with a request for help. Research on proxemic behavior and the effect of status of the person soliciting help provides insight into the possible relationship between status, touch, and compliance. McElroy and Morrow (1994) showed, by manipulating status through the clothing of a male confederate, that an individual of higher status, requesting funds for a charitable organization, was granted more money compared to a situation where he was clothed more informally. However, the funds were significantly higher only when the confederate of high status was placed at about 25/30 centimeters of the subject. When the confederate had a lower status appearance, the reverse effect was observed: being close (30 centimeters) led to a significant decrease in gifts compared to a situation where the physical distance was larger (1 meter).

Thus, research suggests that people tolerate or perceive more positively the invasion of their private space if the invader is of higher status. It is possible that the same occurs with touch. Lynn, Le and Sherwyn (1998) showed that touch of restaurant patrons by a waiter produces a different effect according to the age of the customers. A positive effect of touch on tipping was observed when the patrons were younger than the waiter but a negative effect was found when they were older than the server. In accordance with status norms, older people initiate touch more often than younger people. It seems that transgression of this norm led the "touchee" to more negatively evaluate the "toucher," which, in return, decreased tipping behavior.

In the current experiment the effect of status symbols and touch was evaluated. Insofar as studies show that requests emanating from a high-status person have a higher probability of being accepted, and that there is a positive relationship between proxemic and tactile behaviors, we expected increased compliance to a request to be accompanied by increased status of the solicitor. Because people avoid proximity with lower status individuals, we expected that touch of a low status solicitor would lead to less compliance to a request than when the solicitor was of intermediate status.

METHOD

Subjects

The participants were 90 men and 90 women (age range between 30-50 years), selected randomly in malls of an average-sized provincial city in France. Every fifth participant passing by a definite zone was asked to participate.

Procedure

The experiment took place on sunny days. A confederate, a 27 year-old man, solicited the passersby on the street. The confederate claimed to be conducting an investigation for an association of parents of school pupils and asked the participants if they would agree to answer a questionnaire about television programs for children. The confederate added that the questionnaire included ten questions and that it took 2 to 3 minutes to answer it. According to the condition, the confederate was clothed in an elegant way (suit and tie), conventional (clean jeans and tee-shirt, common standard shoes) or in a very neglected way (worn, dirty and torn trousers, hardly clean tee-shirt, dirty hair). These clothing differences comprised the manipulation of status. A manipulation check of the relevance of this attribution was made among a group of 45 persons taken in the street. They were asked to look at the confederate and to evaluate his social status, using a scale with 9 steps ranging from 1 (lower status) to 9 (higher status). The analysis showed a difference according to the apparel of the confederate: 3.28 for the clothing aspect intended to incarnate the lower status, 5.83 for the conventional level and 7.21 for the elegant appearance. While formulating his request, the confederate, in touch condition, slightly touch the forearm of the person. If the person refused, the investigator thanked him. If he/she complied, then the questionnaire was administered. The questions were of the type: "*Which channel do you think offers the best programs for children?, Do you think that the programs for youth are diversified enough?, What would you like to more often see on the TV as youth programs?, ...*" At the end of the interview, the participants were thanked for their participation.

RESULTS

The dependent variable used in this experiment was the number of respondents who agreed to the request. The results obtained according to gender, status and touch conditions are presented in Table 1 below.

A 2 (touch/no touch) × 3 (high/intermediate/low status) × 2 (male/female) log-linear analysis was used to analyze our data. Results showed a main effect of touch ($\chi^2_{(1,\ 179)} = 7.73$, $p < .01$). When touched, participants complied in 38.9 % of cases, compared to 20.0 % in the no-touch control condition. A main effect of status was found ($\chi^2_{(2, 178)} = 41.93$, $p < .001$) and confirms a linear relation between status and compliance to the request: 58.4 % for higher status, 25.0 % for intermediate status and 5.0 % for lower status. The gender of the participants did not affect compliance ($\chi^2_{(1, 179)} = 0.67$, *ns*). Status by Touch was the only significant interaction ($\chi^2_{(6, 174)} = 14.28$, $p < .03$). However this interaction was qualified by differences between the higher status and intermediate status conditions ($\chi^2_{(4, 116)} = 26.88$, $p < .03$). Indeed, the interaction of Status and Touch in the groups with intermediate status and lower status was not significant ($\chi^2_{(4, 116)} = 8.64$, *ns*). High status compared to low or intermediate status led to increase in the efficiency of touch, whereas low status compared to an intermediate status did not decrease the effect of touch.

TABLE 1 Percentage of Compliance to the Request According to the Conditions of Touch, Status and Participant Gender

	Higher Status	Intermediate status	Lower status
Touch			
Women participants	86.7	33.3	6.7
Men participants	73.3	26.7	6.7
Total	*80.0*	*30.0*	*6.7*
No Touch			
Women participants	46.7	20.0	0.0
Men participants	26.7	20.0	6.7
Total	*36.7*	*20.0*	*3.4*

Note: There were 15 participants in each group

DISCUSSION

Present findings have shown that when the " toucher " was of higher status, there was a significant increase in compliance to the request. This findings confirms the first prediction. Additionally, touch generally increased compliance to a request. These results confirmed those of previous studies (Brockner, Pressman, Cabitt & Moran, 1982; Goldman & Fordyce, 1983; Goldman, Kiyohara & Pfannensteil, 1985; Hornik, 1987; Hornik & Ellis, 1988; Kleinke, 1977a; Smith, Gier & Willis, 1982; Willis & Hamm, 1980). Compared to the intermediate status condition, the high status condition increased compliance to the confederate request, whereas the low-status condition led to a decrease in compliance. Again, such results confirm those of previous studies on helping behavior (Bickman, 1971; Kleinke, 1977b; McElroy & Morrow, 1994).

Contrary to our expectations, touch, compared to non-touch, did not lead to decrease compliance to the request of the low-status confederate. However the absence of an effect must be interpreted with caution. The percentage of compliance to the request was near zero in both levels (touch/no touch) where the status was manipulated. It is possible that the subjects avoided any interaction with the solicitor. Therefore, this avoiding effect does not fully permit the study of the effect of touch on participants' compliance. It is possible that our confederate frightened the participants in the low status conditions, which led them to avoid any contact with him and thus to refuse his request in both conditions.

Surprisingly, no gender effect was found in this experiment. Research had shown that men are more resistant to the effect of physical contact (Fisher, Rytting & Heslin, 1976) especially when the toucher is a male (Roese, Olson, Borenstein, Martin & Shores, 1992). A possible explanation of this absence of gender effect is that the present research was conducted in France, whereas most of the studies cited above were conducted in the United States. Studies have shown that French people are more accustomed to tactile contact than American people (Field, 1999; Jourard, 1966). Perhaps this familiarity led male participants to react positively to tactile contact with a stranger.

Why a stronger effect of touch in combination with higher status? Several explanations are possible. Each of them would require empirical support.

It is possible that touch makes status become more salient and thus mediates a component traditionally associated with status: expertise. Hornik (1987) showed that touch enhanced this perception of competence and professionalism attributed to an investigator. It is thus possible that a reminder of this perception occurred when the " toucher " was of higher status, thus increasing the probability of complying with the request.

It is also possible that being touched by a person of higher status is interpreted as a sign of validation for the self. Howard, Gengler and Jain (1997) found that a sign of respect from a high status person directed toward a low status person enhanced compliance to a request compared to a situation where this sign was not granted. We know that high status people are more likely to initiate touch toward low status people (Hall, 1996) but studies have shown that low status people are more likely to seek tactile contact from high status people (Juni & Brannon, 1981). Perhaps there is a need for implicit recognition among people of lower status. If this sign of recognition is forthcoming, it may predispose the low-status target to comply. This interpretation could explain the results observed by McElroy and Morrow (1994). They showed that the reduced physical distance between a male confederate of high status and a target person when formulating a request led to more compliance than in a comparable condition where the physical distance was larger. However, this effect could not be explained only by the reduction of distance between the subject and the confederate. When the same physical distance manipulation occurred with a requester of intermediate status, the opposite effect was obtained, less compliance with the request.

The explanations proposed here are speculative, and further research is necessary to evaluate the processes or cognitive states activated by tactile contact of a high status solicitor. Personality evaluations of the "toucher" and the self-esteem of the participants must be considered. A *post hoc* control of the participant status could also be made, in order to see if there are differences in touch susceptibility according to the status difference between the "toucher " and the "touchee."

The present results show that it is possible to increase the persuasive effect of touch by manipulating the appearance of the " toucher." Consequently, we might all imagine survey organizations using this knowledge in order to reduce their investigation costs and to increase sample sizes.

REFERENCES

Bickman, L. (1971). The effect of social status on the honesty of others. *The Journal of Social Psychology*, *85*, 87-92.

Brockner, J., Pressman, B., Cabitt, J., & Moran, P. (1982). Nonverbal intimacy, sex, and compliance: A field study. *Journal of Nonverbal Behavior*, *6*, 253-258.

Crusco, A., & Wetzel, C. (1984). The midas touch: The effects of interpersonal touch on restaurant tipping. *Personality and Social Psychology Bulletin*. *10*, 512-517.

Field, T. (1999). Preschoolers in America are touched less and are more aggressive than preschoolers in France. *Early Childhood Development and Care*, *151*, 11-17.

Fisher, J., Rytting, M., & Heslin, R. (1976). Hands touching hands: Affective and evaluative effects on interpersonal touch. *Sociometry*, 39, 416-421.

Goldman, M., & Fordyce, J. (1983). Prosocial behavior as affected by eye contact, touch, and voice expression. *The Journal of Social Psychology*, *121*, 125-129.

Goldman, M., Kiyohara, O., & Pfannensteil, D. (1985). Interpersonal touch, social labeling and the foot-in-the-door effect. *The Journal of Social Psychology*, *152*, 143-147.

Goodman, M., & Gareis, K. (1993). The influence of status in the decision to help. *The Journal of Social Psychology*, *133*, 23-31.

Hall, J. (1996). Touch, status, and gender at professional meetings. *Journal of Nonverbal Behavior*. *20*, 23-44.

Henley, N. (1973). The politics of touch. In P. Brown (*Ed.*). *Radical psychology*. New York : Harper & Row.

Hornik, J. (1987). The effect of touch and gaze upon compliance and interest of interviewees. *The Journal of Social Psychology, 127*, 681-683.

Hornik, J. (1992a). Effects of physical contact on customers' shopping time and behavior. *Marketing Letters, 3*, 49-55.

Hornik, J. (1992b). Tactile stimulation and consumer response. *Journal of Consumer Research, 19*, 449-458.

Hornik, J., & Ellis, S. (1988). Strategies to secure compliance for a mall intercept interview. *Public Opinion Quaterly, 52*, 539-551.

Howard, D., Gengler, C., & Jain A. (1997). The name rememberance effect : a test of alternative explanations. *Journal of Social Behavior and Personality, 12*, 801-810.

Jourard, S. (1966). An exploratory study of body accessiblity. *British Journal of Social and Clinical Psychology*, 5, 221-231.

Juni, S., & Brannon, R. (1981). Interpersonal touching as a function of status and sex. *The Journal of Social Psychology, 114*, 135-136.

Kaufman, D., & Mahoney, J. (1999). The effect of waitresses' touch on alcohol consumption in dyads. *The Journal of Social Psychology*, 139, 261-267.

Kleinke, C. (1977a). Compliance to requests made by gazing and touching experimenters in field settings. *Journal of Experimental Social Psychology, 13*, 218-223.

Kleinke, C. (1977b). Effects of dress on compliance to requests in a field setting. *The Journal of Social Psychology, 101*, 223-240.

Lynn, M., Le, J.-M., & Sherwyn, D. (1998). Reach out and touch your customers. *Cornell Hotel and Restaurant Administration Quarterly, 39*, 60-65.

Major, B., & Heslin, R (1982). Perceptions of cross-sex and same-sex nonreciprocal touch : it is better to give than to receive. *Journal of Nonverbal Behavior, 6*, 148-162.

McElroy, J., & Morrow, P. (1994). Personal space, personal appearance, and personal selling. *Psychological Reports, 74*, 425-426.

Nannberg, J., & Hansen, C. (1994). Post-compliance touch: An incentive for task performance. *The Journal of Social Psychology, 134*, 301-307.

Roese, N., Olson, J., Borenstein, M., Martin, A. & Shores, A. (1992). Same-sex touching behavior: The moderating role of homophobic attitudes. *Journal of Nonverbal Behavior, 16*, 249-259.

Smith, D., Gier, J., & Willis, F. (1982). Interpersonal touch and compliance with a marketing request. *Basic and Applied Social Psychology, 3*, 35-38.

Solomon, H., & Herman, L. (1977). Status symbols and prosocial behavior: The effect of the victim's car on helping. *The Journal of Psychology, 97*, 271-273.

Summerhayes, D., & Suchner, R. (1978). Power implications of touch in male-female relationships. *Sex Roles, 4*, 103-110.

Stephen, R., & Zweigenhaft, R. (1986). The effect on tipping of a waitress touching male and female customers. *The Journal of Social Psychology, 126*, 141-142.

Willis, F., & Hamm, H. (1980). The use of interpersonal touch in securing compliance, *Journal of Nonverbal Behavior, 5*, 49-55.

Do His Friends and Family Like Me? Predictors of Infidelity in Intimate Relationships

Ann Zak, Colleen Coulter, Sabrina Giglio, Jennifer Hall,
Stephanie Sanford
College of Saint Rose

Nancy Pellowski
Grinnell College

While relationship researchers have focused on sexual problems as predictors of infidelity in intimate relationships, we expected that emotional issues would influence unfaithfulness as well. We also believed that problems outside of the relationship itself, in particular a lack of support for the relationship from a partner's family and friends, would predict infidelity. We hypothesized that friend and familial support for the relationship, romantic love and trust would be inversely related to participants' own infidelity. Results supported hypotheses.

Much evidence exists on the factors that predict infidelity (e.g., Roscoe, Kennedy & Cavanaugh, 1989; Shackelford & Buss, 1997). Investigators have emphasized that sexual issues often signal relationship distress and infidelity ahead. For example, Treas and Giesen (2000) found that dissatisfaction with the relationship, strong sexual interest, permissive sexual values and sexual opportunity predict infidelity, while Shackelford and Buss (1997) cited lack of interest in sexual exclusivity and increased or decreased sexual interest as risk factors. Although sexual issues could clearly test relationship exclusivity, emotional issues could also be related to unfaithfulness.

Knox and Gibson (2000), for instance, claimed that falling in love with someone else was the most frequent reason for infidelity among their college-aged sample. Loss of love for a current partner also signals cheating susceptibility (Shackelford & Buss, 1997). Since trust is a strong predictor of both love and satisfaction (Zak, 1999), it could be expected that it, too, is negatively correlated with infidelity. Given that emotional concerns such as love and trust are inversely related to infidelity, issues outside of the dyad but that nevertheless affect the couple could also contribute to partner unfaithfulness. Indeed, Sprecher and Felmlee (1992) found that, among college students, a lack of perceived support for the relationship from partners' families and friends

---

Author info: Correspondence should be sent to: Dr. Ann Zak, Department of Psychology, College of Saint Rose, 432 Western Avenue, Albany, NY 12203, zaka@mail.strose.edu

North American Journal of Psychology, 2002, Vol. 4, No. 2, 287-290.

© NAJP

predicted relationship dissolution both one and two years later. We hypothesized that an inverse relationship exists between a) perceived friend and familial support and infidelity, b) romantic love and infidelity and c) trust and infidelity.

METHOD

Three hundred and forty-five participants were recruited at area colleges and public places in the Northeast and Midwest. Age of participants ranged from 18 to 77 (mean age for women = 26.4 years, *SD* = 1.2, mean age for men = 28.2 years, *SD* = 1.4). Thirty-five percent of the sample was male and 65% was female. Twelve percent of the sample was casually dating, 40% was seriously dating, 8% was engaged, and 40% was married. Participants completed the *Social Support Scale* (Sprecher & Felmlee, 1992), the *Eros* (romantic love) S*ubscale* of the *Love Attitudes Scale* (Hendrick & Hendrick, 1986), the *Trust Scale* (Rempel, Holmes & Zanna, 1985) and the *Infidelity Scale* (Buss & Shackelford, 1997) and were then debriefed. The *Social Support Scale* measures participants' perceptions of friend and familial support for the relationship (e.g., "to what degree do you think that your partner's family approves of this relationship?"), whereas the *Eros Subscale* assesses romantic love (e.g., "my lover and I became emotionally involved rather quickly"). Further, the *Trust Scale* assesses individuals' levels of trust in a partner, with items such as "I can rely on my partner to keep the promises he/she makes," while the *Infidelity Scale* measures one's own unfaithful behavior, with items such as "to what extent are you likely to have a one-night stand with someone besides your partner in the next year?" Cronbach's alpha for the *Social Support Scale* was .82, .76 for the *Eros Scale*, .80 for the *Trust Scale*, and .86 for the *Infidelity Scale*.

RESULTS AND DISCUSSION

In line with past research (e.g., Hendrick & Hendrick, 1986; Sprecher & Felmlee, 1992), we computed correlations separately for males and females. The correlations for men were consistent with hypotheses. For men, negative correlations were found between social support and infidelity, $r_{(121)} = -.36$, $p < .001$, eros and infidelity, $r_{(121)} = -.24$, $p < .01$, and trust and infidelity, $r_{(121)} = -.23$, $p < .01$, whereas for women, inverse relationships between social support and infidelity, $r_{(220)} = -.30$, $p < .001$, eros and infidelity, $r_{(220)} = -.24$, $p < .01$, and trust and infidelity, $r(220) = -.46$, $p < .001$ were shown (see Table 1). It is interesting that the correlation for women between infidelity and trust ($r = -.46$) is higher than the correlation for men between infidelity and trust ($r = -.23$). Although the correlations are not significantly different from one another, it appears for women that trust is somewhat more important, especially as it relates to one's own cheating behavior. It seems as if females who do not trust their partners are more likely to contemplate committing infidelity than men who do not trust their partners.

TABLE 1 Correlates of Infidelity for Both Sexes

	Men Infidelity	Women Infidelity
Social Support	-.36**	-.30**
Eros	-.24*	-.24*
Trust	-.23*	-.46**

*p <.01 **p <.001

TABLE 2 Correlates for all Scales

	Social Support	Eros	Trust	Infidelity
Social Support		.62**	.68**	-.28**
Eros			.74**	-.22**
Trust				-.34**

**p <.001

Overall, the findings indicate that persons who experience social support for their relationship, romantic love and trust are less likely to commit infidelity. With 38% of college students reporting infidelity in their romantic relationships (Knox & Gibson, 1997), unfaithfulness is a widespread problem. Future researchers may wish to examine sex differences and other "buffers" against infidelity.

REFERENCES

Buss, D. M. & Shackelford, T.K. (1997). Susceptibility to infidelity in the first year of marriage. *Journal of Research in Personality, 31,* 193-221.

Hendrick, C. & Hendrick, S.S. (1986). A theory and method of love. *Journal of Social and Personal Relationships, 5,* 161-183.

Knox, D., & Gibson, L. (1997). Why college students end relationships. *College Student Journal, 31,* 449-453.

Rempel, J.K., Holmes, J.G. & Zanna, M. (1985). Trust in close relationships. *Journal of Personality and Social Psychology, 49,* 95-112.

Roscoe, B., Cavanaugh, L. E. & Kennedy, D.R. (1989). Dating infidelity: behaviors, reasons and consequences. *Adolescence, 23,* 35-43.

Shackelford, T.K. & Buss, D.M.(1997). Cues to infidelity. *Personality and Social Psychology Bulletin, 23,* 1034-1046.

Sprecher, S. & Felmlee, P. (1992). The influence of parents and friends on the quality and stability of romantic relationships: a three wave investigation. *Journal of Marriage and the Family, 54,* 888-901.

Treas, J. & Giesen, D. (2000). Sexual infidelity among married and cohabitating Americans. *Journal of Marriage and the Family, 62,* 48-61.

Zak, A. (1999). *Love on the rocks: influences on trust and satisfaction in intimate relationships.* Paper presented at the Eastern Psychological Association Conference, April, 1999.

An Interview with Mark Ylvisaker About Students with Traumatic Brain Injury

Mark Ylvisaker
College of Saint Rose

(interviewed on behalf of NAJP by)
Michael F. Shaughnessy
Eastern New Mexico University
(and)
Dan Greathouse
Portales School System

Biographical Sketch - After an early career teaching philosophy at Carleton College and the University of Pittsburgh, Dr. Ylvisaker became a speech-language pathologist. Currently, he is Associate Professor of Communication Disorders at the College of Saint Rose in Albany, New York. He has 25 years of clinical and program development experience with children and young adults with neurogenic cognitive, behavioral, and communication disorders in rehabilitation and special education settings. Dr. Ylvisaker has approximately 100 publications on brain injury-related topics, including the recently published books *Traumatic Brain Injury Rehabilitation: Children and Adolescents* and *Collaborative Brain Injury Intervention: Positive Everyday Routines*. He is actively engaged in the work of several professional organizations, serves on a number of national and international committees and editorial boards, consults to rehabilitation programs in 11 countries on five continents, and presents frequently to national and international audiences on topics related to rehabilitation and education of individuals with cognitive, behavioral, and communication disability.

NAJP: What are you currently working on or researching?

MY: My main current interest is in the process of creating and maintaining well conceived and flexible supports for children, adolescents, and young adults with persistent impairment after brain injury. In the case of children and adolescents, that means working collaboratively with parents and school staff, including assistant level staff, to identify the most useful ways to interact with, teach, and

---

Author info: Correspondence should be sent to: Dr. Michael Shaughnessy, Psychology Dept., Eastern New Mexico University, Portales, NM 88130.

North American Journal of Psychology, 2002, Vol. 4, No. 2, 291-308.
© NAJP

otherwise support the child. Then we modify those supports over time as the child, the environment, and the demands of the academic and social curricula change. To this end, my colleague Tim Feeney and I use and try to teach support people the process of ongoing contextualized, collaborative, hypothesis-testing assessment. Our goal is not simply to create an intervention and support plan that meets current needs. More importantly, we attempt to leave behind the expertise and optimism needed to revise that plan in an ongoing way over the student's developmental years. We are attempting to systematize this consultation process and then evaluate its effectiveness in a systematic manner.

NAJP: Have we seen an increase in pediatric TBI and awareness of TBI?

MY: These are two quite separate questions. The epidemiologic question is a tough one – and in dispute. For example, in 1997-98, 11,895 students were reported to the US Office of Special Education Programs as having a TBI classification, versus 2.7 million with a specific learning disabilities classification (250:1 – LD:TBI) out of a total special education population of 5.4 million students. With such small reported numbers, it is understandable that relatively little attention is paid to TBI in special education funding, teacher training, and the like. On the other hand, my colleagues and I have estimated that approximately 20,000 new cases of school age (or younger) TBI-related persistent disability are added annually, bringing the prevalence total in schools in the US to a number many times larger than that reported to the government.

There is a suggestion in the epidemiologic literature that prevention efforts have lowered the incidence of pediatric TBI. However, this suggestion may be an artifact of changes in hospital admission criteria rather than a true incidence reduction.

With respect to awareness, most state departments of education now have some person or office responsible for coordinating information about TBI, many states have developed TBI manuals for educators, and the number of training sessions for educators seems to have increased. Having said that, I must add that when I consult in schools, I rarely encounter education staffers who have received specific training in TBI.

NAJP: In 1990, TBI was added to the Individuals with Disability Education Act as an educational disability. Why was this important?

MY: During the 1980s, I was one of the many professionals in the field who believed that, on balance, it would be helpful to add TBI as an educational disability. There were two primary reasons: First, many of the students with TBI with whom I worked at that time in Pennsylvania had difficulty qualifying for special education and therapy services when they returned to school. Often their profiles on testing failed to match profiles of educational disability written into special education regulations. For example, many of the students with even fairly severe injuries ultimately scored within normal limits on IQ tests (although perhaps considerably lower than pre-injury performance) and reasonably close to grade level on educational tests. As you know, IQ tests are notoriously weak at identifying the effects of the most common kinds of injury in closed head injury, namely prefrontal and limbic system damage. Furthermore, the kinds of academic knowledge and skill measured by most tests are associated with posterior parts of the brain, which are relatively less vulnerable than anterior parts. Thus, over and over again, we had to work very hard to acquire needed services and supports for students whose test profiles at the time of their return to school did not seem to indicate the need for such services.

Second, back in the 1980s we expected that federal recognition of TBI as an educational disability would trigger increased recognition of this population, more support for pre-service and in-service training of education professionals, and increasingly well conceived educational programs for the students. To a limited extent, this expectation has been realized.

I must say, however, that I was cautious in my advocacy for federal recognition of TBI as an educational disability. My fear was that such recognition would inevitably be followed by development of "TBI assessment batteries" and "TBI Educational Curricula" – and unfortunately these fears have to some extent been confirmed.

NAJP: What exactly is wrong with "TBI assessment batteries" and "TBI Educational Curricula?"

MY: First, there is far too much diversity within the population of students with TBI to justify population-specific tests and curricula. The severity spectrum is enormous – ranging from complete recovery following very mild injuries to massive and persistent impairment. And virtually any combination of functions can be spared or impaired following head trauma.

Second, there is great overlap in educational and therapy needs between the population of students with traumatic, or more generally acquired brain injury, and other special education populations. First, many students with TBI had some type of pre-injury educational disability that did not disappear with the injury. In addition, selective executive system impairment, common in TBI, is shared with other populations, such as ADHD, Tourette's Syndrome, and some forms of autism. Furthermore, there are many validated approaches to learning and processing difficulties that are shared by many categories of special education students. Therefore, there is every reason to make assessment, intervention, and support decisions based on the specific student's presenting profile of abilities and needs rather than on the etiologic label, TBI.

In this connection, I think it's worth remembering the last major attempt to create specific educational prescriptions for children labeled "brain injured." In the 1940s and 50s, there was a substantial literature on "brain-injured students," motivated largely by Alfred Strauss, whose inspiration came from Kurt Goldstein's work with soldiers returning from World War I with acquired brain injuries. Strauss saw parallels between the brain injury symptoms described by Goldstein and characteristics of children with mental retardation or other poorly defined mental health conditions (e.g., inattention, concreteness and rigidity in thinking, inefficient concept formation, poorly regulated behavior). Increasingly he and his colleagues applied the term "brain injury" to children with no documented history of insult to the brain. In recent decades, this "Strauss syndrome" has evolved into more precise diagnostic categories, including ADHD and specific learning disabilities.

More importantly, Strauss and his colleagues (e.g., Newell Kephart and Laura Lehtinen) set about to create intervention models that they believed should be applied to all of the children who were captured under their broad "brain injury" net. The model included rigid environmental prescriptions (e.g., an extreme reduction in environmental stimulation) and basic-level cognitive and perceptual-motor training exercises. A half-century of educational research, from the 1950s to today, has demonstrated that there was relatively little therapeutic or educational merit in the intervention plans prescribed within the Strauss framework. (See Lester Mann's painstakingly thorough history of process training, *On the Trail of Process*.) However, those intervention prescriptions of the 40s and 50s bear an eerie and sobering similarity to some of the educational recommendations derived from the neuropsychological framework of the 80s and 90s, and applied to many students with TBI today.

So, for these historical and conceptual reasons, I take every opportunity I am given to caution professionals against anything resembling a "TBI assessment battery" or a "TBI educational curriculum."

NAJP: Many specialists have promoted the use of programmed software to "rehabilitate" attention, memory and problem solving. Your comments about the relative ineffectiveness of perceptual-motor and cognitive training exercises of 30 and 40 years ago suggest that you might be skeptical about today's technologically more sophisticated cognitive training exercises.

MY: Exactly. I think that the research with adults with TBI has largely supported such skepticism. For many years, researchers have been in general agreement that memory cannot be "rebuilt" with decontextualized remembering exercises, although there certainly are many useful personal strategies, teaching strategies, and environmental compensations that can be useful for people with memory impairment. More recently, reviews of the research literature on attention training have suggested that, when the studies are well controlled, the effectiveness of such training is very limited. And as a former teacher of logic – and as a person who tries to keep up with the literature on cognitive theory – I think I can say with some confidence that abstract, decontextualized exercises in problem solving or reasoning are poorly conceived, if only because there is such a fierce domain specificity to skill in these areas of cognition. For sure, people can become better thinkers, reasoners, and problem solvers, but it makes all the sense in the world to practice these skills within specific academic, social, or vocational domains in which the skills are needed.

Robert Sternberg, who was featured in one of your earlier interviews, has suggested that four conditions must be met for transfer of training to be likely: encoding specificity (Tulving's term), adequate organization of the to-be-learned information or skill, discrimination ability, and a psychological set that predisposes the person to seek transfer opportunities. Taken together, these conditions seem to argue for a highly contextualized approach to cognitive intervention. And of course such an approach is also intimately associated with the large family of cognitive theories that fall under the labels situated cognition, apprenticeship learning, and Vygotskyan psychology.

Finally, it seems to me that contextualized intervention meets the gold standard of plain old fashioned common sense. For example, if psychology interns or student teachers are poor reasoners or problem solvers in their chosen professions, we don't send them off to a logic class for reasoning exercises. Rather, we try very hard to teach them how to organize their thinking within the specific domains in which they have to think in an organized manner. I think the same principle applies to students with disabilities.

NAJP: Given your comments about variety in the population and your cautions about popular cognitive retraining programs, I take it that you would be reluctant to talk about something as general as "state of the art" TBI treatment?

MY: First let's make clear that we are not talking about medical management, pharmacological management, or physical restoration. "State of the art" management in those domains would be a very different conversation. I agree that state of the art intervention at the level of specific educational curricula or intervention plans is a misleading and distracting goal. However, at a more abstract level, I think we can talk about state of the art intervention planning, based on flexible, contextualized hypothesis testing, and state of the art intervention for students with cognitive and behavioral impairment that highlights the use of customized and flexibly applied antecedent supports.

Having made the case for diversity, flexibility, and an experimental approach to intervention planning, let me add that there are some common themes that run through the population of students with TBI – at the level of impairments, educational needs, and intervention plans. Many of these themes are associated with vulnerable prefrontal and limbic system structures.

NAJP: Yes, we were going to ask about that. Your writings have emphasized specific approaches to executive system impairment and behavior management.

MY: Let's start with executive system impairment. The neuropsychology literature has many lists of functions said to comprise the executive system, some factor analytic descriptions, and a small number of attempts to capture the executive system simply as deliberate goal-oriented behavior. For example, Muriel Lezak has characterized the executive system as the processes needed to formulate goals, plan ways to achieve them, and execute the plans. Over the years, I have used the following operational definition, largely because it is consistent with the literature and lends itself to functional intervention: executive functions include self-awareness of strengths and limitations, ability to set reasonable goals, ability to plan and organize behavior in pursuit of the goals, ability to initiate goal-directed behavior and inhibit behavior that interferes, ability to monitor ongoing behavior and evaluate it in relation to achievement of goals, and ability to flexibly shift sets and revise strategies in the face of obstacles or failure. All of these abilities (1) come in degrees, (2) have a developmental trajectory that starts early in life and progresses slowly and in interaction with other developing abilities, and (3) vary in relation to context and content.

Clearly some aspect of executive functioning can be relatively impaired and others relatively spared following brain injury. For example, ventral prefrontal injury appears to differentially affect inhibition whereas dorsomesial prefrontal injury appears to differentially affect initiation. However, it is possible and I think desirable to target aspects of executive functioning within the integrated context of functional goal-directed behavior that includes all of the listed dimensions. Avoiding the "disassembly" of the executive system in intervention has the advantage of avoiding the massive transfer problems associated with all cognitive training approaches that disassemble the mechanism for purposes of applying decontextualized retraining exercises.

There was a time in the not-too-distant past when neuropsychologists failed to appreciate the fact that executive functions begin their developmental trajectory in infancy and that early experiences influence their long-term development. More recently, however, early and protracted development of executive functions has been highlighted by several investigators of early child development, including Bruce Pennington and Martha Bronson. Recognizing the early development of executive functions and their complex interaction with other aspects of cognitive and social development leads directly to an intervention focus on the executive system as early as preschool and also early in recovery after brain injury.

We use the term "executive system routines" or "self-regulation routines" in working with parents and school staff. The point is to put in place in the child's life routines of executive self-regulation with as much support as the child needs to be successful and with adult scripts designed to be internalized over time by the child in the process of creating internal self-regulation. In schools, we typically work with staff to identify executive function support routines that best fit the child's current functioning, and then video tape these routines in the school setting so that others can see concretely what these routines of interaction look like.

NAJP: Can you give an example of an "executive function routine"?

MY: Sure. I recently worked with school staff serving a kindergartner who had sustained a TBI at age two. As is so often the case, her behavioral difficulties had escalated as developmental tasks became more difficult by age five and then worsened as she started kindergarten. Staff had begun to talk about referring her to a special school.

At the beginning of kindergarten, staff would present her with a task; she would indicate that she could do it; but then when she realized that it was hard for her, she would run out of the room or act out in some other way. She was becoming very manipulative. So we put in place everyday school routines that included the following elements: (1) staff were very clear in their presentation of tasks and ensured that they were within the girl's capacity; (2) staff would generally either ask the girl if she thought the task would be easy or hard – or simply state their own view of the task's difficulty level; (3) in the case of difficult tasks, staff had a brief chat with the girl about what could be done to succeed despite the difficulty; (4) staff then provided needed support, but only in those areas of the girl's known weakness; (5) finally, staff ensured a review of all completed tasks, highlighting the difficulty level in relation to the girl's abilities and what they did to succeed.

Over the year, the girl took over more and more of the responsibility for identifying the tasks difficulty level, the plan for success, and the review. This is just one example of an EF routine – but by the end of the school year, after hundreds of these highly contextualized routine interactions, the girl was generally accurate in identifying tasks as hard or easy for her, in asking for help if necessary, and in knowing what good help looks like. In addition, her refusing to work and fleeing down the hall were a distant memory.

In May, we video taped these classroom-based EF routines for the fall staff (and included commentary by the kindergarten teacher), because the girl was scheduled to change schools. By report she is doing well this year in a mainstream setting, in part because the staff know their scripts and their role in facilitating development of executive functions.

NAJP: And what about behavior management?

MY: Because of vulnerability of self-regulation centers in the frontal lobes and because pre-injury behavior problems predispose some children to injuries, behavioral issues are common after TBI. Furthermore, the behavioral concerns create the toughest problems for parents and teachers alike. Most schools, and I think most parents, organize their behavior management strategies around consequences. That is, when the child acts in a desirable manner, adults offer some kind of reward; and when the child acts in an undesirable manner, adults either try to ignore the behavior, remove the child from ongoing activity ("time out"), or impose a punishment. Contingency management of this sort has a long and in many cases noble history.

However, children and adolescents with TBI often respond inefficiently at best to these common behavior plans. There are at least five possible explanations for this inefficiency. Antonio Damasio has repeatedly suggested that ventral prefrontal areas of the brain are critical for connecting memories of what we have done (conscious or unconscious) with our gut-level reactions to these events ("somatic markers"). When these connections are not made, rewards and punishments cease to be effective in guiding future behavior. Others have highlighted the role of inhibition impairment (also associated with vulnerable ventral prefrontal systems) in reducing the effectiveness of contingency behavior management. This makes sense for the same reason that one does not try to control the potentially disastrous effects of exploratory toddler behavior by reacting to the disasters, but rather by preventing the disasters by child proofing the house. Others have highlighted impaired working memory, impaired initiation, or simple oppositionality as explanations for the relative inefficiency of contingency contracts for students with behavior problems after TBI.

Whatever the explanation, many students with TBI are ideally suited for behavior plans that emphasize antecedent control procedures. Over the past decade, antecedent-focused behavior management has come to be organized under the heading Positive Behavior Supports, which has been investigated experimentally for the most part with children and adults with developmental disabilities. Tim Feeney and I have published a few papers in which we used single subject experiments or less systematic case material to illustrate application of antecedent control procedures to students with behavior problems after TBI. The general goal is to put proactive supports in place so that (1) negative behavior based on impulsiveness is infrequently triggered, (2) negative behavior becomes inefficient and unnecessary for the student in everyday routines, and (3) because of the antecedent supports, the student develops habits of behavior that are positive and generally satisfying.

Of course there is a good deal of procedural detail here and individual support plans need to be designed using the contextualized, collaborative, hypothesis-testing process that I mentioned earlier. Tim and I recently published a paper in the journal *Educational Psychology in Scotland* in which we argued that positive behavior supports in behavioral psychology and Vygotskyian apprenticeship approaches to cognitive impairment are comfortable theoretical bed fellows and yield procedural details that are very similar.

NAJP: Are you saying that consequences are not relevant to behavior management of these students?

MY: Not entirely, no. I think that consequences are relevant for the following reasons. First, like all neurologic impairment, inefficient response to consequence management is relative. Furthermore some people with TBI entirely escape injury to those parts of the brain related to effectiveness of traditional behavior management. Second, all children need to learn how the world works. So consequences that are naturally and logically related to negative behavior may be important even if those consequences cannot be expected to modify the behavior. Third, some people need the threat of a major negative consequence to motivate them to take antecedent-focused plans seriously. This is true of many of the young adults with whom we work. A potential consequence, like prison for example, may bring them to the table, so to speak, in developing everyday routines that feature prevention and other antecedent supports. However, schools rarely have anything resembling such a major threat to bring tough kids to the table. And importantly, in these cases it is not the consequences that modify the behavior. Rather a credible threat of consequences helps to underscore the importance of the positive antecedent-focused routines.

So consequences must be thoughtfully included in behavior plans. However, in most cases it's the positive habits created by intelligent manipulation of antecedents that has the desired effect.

NAJP: You have emphasized the usefulness of organized, customized hypothesis testing. What about the formal assessments that are typically conducted by neuropsychologists in hospitals and the testing that schools may require for acquisition of IEP-related services?

MY: First, I need to emphasize that in discussing hypothesis testing, I was addressing assessment for one purpose only, namely designing the most effective package of interventions. There are, of course, many other purposes that assessments can serve. In relation to educational planning, neuropsychological assessments completed during inpatient admissions face three large obstacles. First, children are typically still undergoing neurologic recovery at that time, so when they reenter school the neuropsychological report may already be out of date. Second, most hospitals lack realistic school-like settings and routines that neuropsychologists can use to test the ecological validity of their office-bound findings. Third, even if school staff receive the reports and have time to read them (which is often not the case), the professional language common to medically-oriented reports may be foreign to them.

Formal assessments conducted by school psychologists early after reentry can be useful. However, there is still the issue of ongoing spontaneous recovery. Furthermore, school psychologists often lack the neuropsychological training needed to plan an assessment for a student with an unusual neurological profile. And, as I mentioned earlier, commonly used intelligence and psycho-educational tests may yield misleading findings for children with frontal lobe injury.

NAJP: So, what do you recommend?

MY: I think that neuropsychological assessments can be very helpful if two conditions are met. First the neuropsychologist needs to understand child development, the school context, and the needs and language of school staff. Second, the report should outline and recommend school-based exploration of the themes that emerge from neuropsycholgical assessment, including intervention trials and diagnostic teaching that school staff can implement in the school setting.

If the referral for neuropsychological assessment occurs after the student has returned to school, that assessment will be useful to the extent that the neuropsychologist is willing to interact with school staff in planning the assessment and also in exploring practical implications of the findings. I make these comments based on hundreds of consultations in schools in which these conditions for neuropsychological assessment were not met and consequently the resulting report played no role whatsoever in helping staff develop the educational program.

NAJP: How soon following the injury should schools be responsible for conducting formal education evaluations?

MY: Again, questions about assessment must always be answered in relation to the purpose of the assessment and the validity of the assessment tools in relation to that purpose. In my judgment, extensive office-bound testing is largely a waste of time when the student remains in a period of relatively rapid neurological recovery, assuming that the testing is for the purpose of designing an educational program. However, quick tests designed to reveal levels of functioning in basic academic areas (e.g., reading, writing, math) may be useful for teachers.

On the other hand, if the goal is to ensure that the student qualifies for special education services and the student's school district requires test results, then the minimum required testing may need to be done. However, even with qualification for services as the goal, testing may be contraindicated if the student's profile of impairment is known to render certain assessment tools invalid (e.g., selective prefrontal injury in relation to intelligence tests).

NAJP: In some of your work you have advised schools to create temporary IEPs for students with brain injuries because of rapid changes. What is your practical advice for how often IEPs should be modified? How often should formal re-evaluations be conducted?

MY: This varies with the severity of the injury, rate of recovery, and other issues. We have recommended a red flag/safety net system for students returning to school after relatively mild injuries. Briefly, the idea here is that somebody at the hospital (e.g., emergency room nurse; discharge planner) should alert the school (with parents' permission) to the possibility that the student may need temporary accommodations if symptoms emerge following return to school (i.e., "red flags"). We list common red flags in this protocol and commonly used temporary accommodations. This protocol, published in the *Journal of Head Trauma Rehabilitation* in 1995, can be implemented in a way that is very simple and circumvents the expensive and cumbersome IEP process.

In the event of more severe injuries, ongoing change and potentially unpredictable reactions to being back in school may indicate IEP reviews every 2 or 3 months over the first year – but this has to be an individualized decision. And these frequent reviews rarely involve retesting. Formal re-evaluations, including extensive testing, are an entirely different question. Again, if the goal is to refine the intervention program, then testing should be used only if the tests can be expected to answer questions related to educational planning. And generally such tests would not be needed for this purpose more frequently than once a year.

NAJP: Because children with brain injury may continue to improve neurologically and also demonstrate inconsistent behavior day-to-day, is there a need for the development of a different IEP approach regarding TBI students? If so, what type of approach would you suggest?

MY: I don't think it's wise to *over*-emphasize the differences of students with brain injury, if only because that emphasis may intimidate teachers and result in their failing to apply the basic competencies and insights they possess as educators. But you are right about the possibility of change and inconsistency in performance. Perhaps the main IEP difference is the need to emphasize ongoing contextualized hypothesis-testing assessment in the IEP. Some educators and administrators believe that an assessment should be completed quickly when the child returns to school, followed by development of an IEP that includes placement, goals and objectives, and procedures for achieving the objectives – but no more assessment until triggered by a periodic IEP review. However, students who can be expected to change and who have confusing profiles of ability and disability require ongoing exploration of their learning and self-regulatory capacities and of teaching and management procedures. Thus IEP goals should include this ongoing flexible assessment process.

NAJP: What do you see as the single most overwhelming problem with teaching the TBI student, as service delivery models shift to serving students in inclusive settings?

MY: I think that the answer to this question depends a lot on age and educational level. For example, in preschool classrooms and in the early grades, individualized teaching, peer support, provision of related services, and customized use of assistants can often be organized into the routine of an inclusive classroom with little need for pull-out services (except perhaps to experiment with routines, supports, and teaching styles). In the later grades and especially in middle school and high school, integration of special services and supports in an inclusive environment becomes more challenging. Furthermore, students at this level may be extremely resistant to being stigmatized with anything "special" in a classroom with regular education peers. In my experience, there is a very important role for special education resources rooms for these students. In the resource room, the student can get organized, explore the usefulness of special strategies, receive intensive tutoring in areas of significant need, and possibly ventilate or receive important emotional support. This need for resource room time has also been born out by studies of strategy intervention with students with leaning disabilities.

The other major theme, in my experience, is training and support for educational assistants. In the case of students with significant impairment, assistants may be the most critical members of the instructional team, but they receive the least training and support. Therefore, when Tim and I consult in schools, we often focus much of our time on assistants, creating and modifying routines of support so that the assistant knows how to help without creating helplessness and without generating opposition from the student.

NAJP: You have done extensive research regarding school reentry. However, should there be a simple, standardized protocol developed for schools to ensure successful exchange of information regarding educational implications for the TBI student? What is available on the market today? Is the case management checklist mentioned in your work *School Reentry Following Head Injury: Managing the Transition from Hospital to School* recommended?

MY: I hate to sound like a stuck record, but it remains critical to appreciate the variety within this population and therefore the need for flexibility in all aspects of the student's management. The checklist you mentioned comes from a 10-year old publication based on our work in pediatric rehabilitation in the 1980s. I think that there continues to be great value in that checklist for children whose inpatient rehabilitation admission continues for several weeks or perhaps months. In these cases, rehabilitation staff can and should create school-like environments, explore school-related variables listed in that checklist, and ideally prepare a video tape that shows concretely what works and what does not work for the student. However, with dramatically reduced lengths of stay over the past 10 to 15 years in children's hospitals and rehabilitation hospitals, much of this hospital-based work is unrealistic today. Thus greater responsibility may fall on the school to engage in relevant exploration of variables that affect learning and self-regulation. In more recent publications, Tim and I have proposed checklists for program planning tied to specific profiles of strength and weakness, especially profiles that may be confusing to school staff. (See our chapter on school reentry in my 1998 book, *Traumatic Brain Injury Rehabilitation: Children and Adolescents*.)

It is also important to remember that year-to-year transitions within the student's school career may be as important as the hospital-to-school transition. This is particularly true of students who "grow into their disability," that is, those students whose disability increases over time, perhaps in part because their injury included parts of the brain that need to continue to develop neurologically in order to support later developmental transitions. In these cases, we recommend an end of the year video tape to show next year's staff what works and what does not work. This can often be done with relative ease with preschoolers and grade-school age children. In secondary school, year-to-year staff transition conferences may be the best mechanism.

NAJP: How does the school district best prepare the school personnel regarding the needs of TBI children?

Here again there are several possibilities. During the hospital-to-school transition, it is often possible for hospital staff to visit the school, address the relevant school staff, and answer their questions about the implications of the injury. Some states or regional education agencies have full-time TBI consultants or teams of peer consultants who can spend time orienting staff who are unfamiliar with the population. I think that the two most effective procedures for helping prepare staff for their work with a student with TBI are (1) actively engaging that staff person in contextualized assessment experiments so that they learn while doing, with guidance, and (2) showing illustrative video tapes of the student engaged in relevant routines so new staff can see what good teaching, interaction, and support look like.

NAJP: Do you have any substantial suggestions for handling 5 to 6 hour IEP meetings which are focused around the TBI student? Is there a successful way to break these meetings up into smaller meetings? Any suggestions?

MY: Let's try to break the topic down into parts. First, there may need to be a general "TBI 101" session and general orientation to the student's impairment and needs at the time of reentry. This is probably a half-day commitment, but is separate from the IEP process. The first (very general) IEP may be written while the student is still in the hospital, giving school staff exploration time before they are responsible for their first IEP. Whether developed before or after reentry, the first IEP should be honest in its inclusion of many ongoing assessment goals (which we discussed earlier). This reduces the time needed for the first IEP meeting.

It seems to me that the development of subsequent IEPs is ideally a task spread over time (as opposed to being concentrated in one meeting) and one that includes participation of the student. In the chapter on executive functions in my 1998 pediatric TBI book, we included a student self-assessment form that can serve as a critical part of the IEP development process and as an important component of EF intervention at the same time. That chapter also includes an IEP development flow chart that includes initial presentation of the IEP to parents as separate from the formal IEP meeting. If these steps are followed, then the legal IEP meeting is short and efficient, because the bulk of the work has been done in advance as part of ongoing intervention.

NAJP: How does the school district prepare peers who may look forward to their friend returning to school, only to find he or she is not the same person?

MY: If the hospital stay is protracted, somebody should organize a system for two-way communication between peers in school and the hospitalized child. If this communication is by video tape, then there may be less shock when the student returns. It is also desirable to gather peers together to offer them information and guidance before the child returns. This may be done by a hospital professional, but may be more effective if the presenter is a trusted school person who has technical information provided by hospital staff. Finally, some children with brain injury are willing to talk to their peers about their experience and possibly even describe some of their difficulties and how others can help. It is important to organize such presentations so that the student with a brain injury is presented as a returning hero rather than a supplicant for understanding and mercy.

NAJP: Thinking of intervention as revolving around the everyday routines of life would seem to make parents key players in the rehabilitation of these children. What kind of counseling and education should parents receive?

MY: You're absolutely right about parents being key players. It seems to be simple common sense that the adjustment and supportiveness of parents would have an impact on the child's outcome. But this point has also been confirmed by careful studies conducted by the Ohio research group that includes Gerry Taylor, Dennis Drotar, Keith Yeates, and Shari Wade. More importantly, a large randomized clinical trial in Brazil conducted by Lucia Braga showed that rehabilitation efforts delivered by specialists largely working indirectly through well oriented and trained parents were more effective than services delivered largely by specialists directly treating children with brain injury in a clinical setting.

Of course there is as much variety among parents of children with TBI as among the children. So education and counseling needs vary and must be individually adjusted. In my judgment, the most effective education is offered by specialists rolling up their sleeves and working elbow to elbow with parents in identifying the most effective intervention strategies and engaging parents in individualized problem-solving exercises focused on their child's specific needs. Currently Shari Wade is systematically investigating the effectiveness of a problem-solving approach to parent education and support. Her intervention model is a modification of the problem-solving approach of Nezu and D'Zurilla, which has been validated with other clinical populations.

Over the long haul, I think that parents do best when they understand life with a child with chronic impairment after brain injury as an ongoing juggling act – and the items being juggled are themselves balancing acts. This mixed metaphor -- effectively juggling balancing acts – seems to describe the experience of successful families. Among the balancing acts that need to be kept in the air at the same time are the following: (1) Balancing alertness to ongoing developmental challenges with optimism that problems can be overcome and that the long-term outcome will be OK; (2) balancing the natural desire to protect the child with the need to provide the child with normal experiences allowing for developmental growth; (3) balancing the need to impose adult authority with sufficient opportunities for child decision making; (4) balancing a desirable consistency in the child's life with the need for ongoing experimentation to determine what routines and supports work best; (5) balancing a natural focus on the child's special needs with a commitment to normal childhood experiences; (6) balancing respect for the expertise of professionals with recognition of the fact that ultimately it's the everyday people in the child's life who have the greatest impact; (7) and in the case of children with organically based behavior problems, balancing an understanding of the child's self-regulatory difficulties and need for support with a system of accountability designed to teach the child that some behavior may be understandable, but is ultimately inexcusable.

Each of these balancing acts is worth a separate discussion. My point right now is that most parents somehow need to gain reasonable comfort with these balancing acts. Parents who are further along in their adjustment process are often the most effective source of inspiration and support for parents who are struggling.

NAJP: Which of your two books, *Collaborative Brain Injury Interventions: Positive Everyday Routines* or *Traumatic Brain Injury Rehabilitation: Children and Adolescents*, would be the more appropriate resource for public schools?

MY: Thanks for your interest. Both books were published in 1998 and have information relevant to teaching and supporting students with brain injury. I would guess that the children and adolescents book is more directly related to the interests of educators and related services providers.

NAJP: What is your professional opinion about alternative therapies, such as hyperbaric oxygen and various types of neuro-feedback therapies for the treatment of traumatic brain injury?

MY: I can't speak as an expert about these therapies, but I do think that families should be urged to do their homework before committing themselves to therapies that often carry a substantial price tag – often at enormous expense. I have known families who have exhausted their resources on such treatments with little functional improvement in their loved one.

Michael F. Shaughnessy is currently Professor of Psychology at Eastern New Mexico University and he is a member of the New Mexico Traumatic Brain Injury Advisory Board. Dan Greathouse is an Educational Diagnostician with the Portales School System and he works with students with open and closed head injury.

Delivering Human Services to Native Americans with Disabilities: Cultural Variables & Service Recommendations

Deborah Rowley & Ruth Anne Rehfeldt
Southern Illinois University

Providers of social services in the United States have long recognized the need for services by members of Native American populations, but have struggled with the differences between American Indian and non-Indian culture that seem to impede effective service delivery. This report discusses the causes and prevalence of disabilities among Native American communities, and elaborates upon aspects of Native American culture that may affect how social services are delivered and received. Recommendations are also made for service providers to engage in culturally-sensitive practice.

Human service providers in the United States have long been challenged by the difficulties encountered when providing services to members of Native American or American Indian populations. Service providers, be they social workers, case managers, rehabilitation counselors, or home health workers, are often unprepared for the vast differences in cultural backgrounds that separate them from their Native American consumers. Moreover, such personnel are often unaware that their service delivery practices may conflict with the beliefs and customs that are at the very core of Native American culture. The purposes of the present paper are threefold: First, we will explore the ways in which disability and disease are conceptualized by many Native American groups, so that human service providers can be informed of the impact of these belief systems on services. Second, we will explore the existing barriers to effective service delivery to members of Native American populations; and third, we will offer recommendations for reducing or eliminating the impact of those barriers so that members of Native American communities can benefit from needed social services.

Disability Prevalence and Causes

The prevalence and severity of disability is greater for American Indians than it is for non-Indians. This includes conventional developmental disabilities such as mental retardation, cerebral palsy, and sensory impairments, as well as fetal alcohol syndrome and fetal alcohol

---

Author info: Correspondence should be sent to: Ruth Anne Rehfeldt, Ph. D., Rehabilitation Institute, Mailcode 4609, Southern Illinois University, Carbondale, IL 62901.

North American Journal of Psychology, 2002, Vol. 4, No. 2, 309-316.

© NAJP

effect. In fact, the latter disorders were recently reported to be the most prevalent type of disability among Native American communities, and Native American infants have twice as much risk for being born with fetal alcohol syndrome or fetal alcohol effects as non-Indian infants (Morgan, Guy, Lee, & Cellini, 1986). Hence, substance abuse is also a common disability or disorder in such communities, as are physical disabilities resulting from alcohol-related automobile accidents.

Factors contributing to such high use of alcohol include poverty, unemployment, illiteracy, and depression (Kaltenbach, 1980). According to the 1990 Census Bureau, 22.1% of Native American families in Chicago live at or below the poverty level, with an annual income of $11,251 or less. A single woman heads 49.7% of these families. Native Americans have a 70% high school dropout rate (Census Bureau, 1990). With no education, individuals are at greater risk for substance abuse. Living at the poverty level makes malnourishment likely, which may play a role in the development of numerous disease and disorders. Other factors which lead to increased risk of disability are substandard housing, environmental toxins, high stress, and neighborhoods with high crime rates. Thus, there are many causes of disability in Native American communities. Not only must services for consumers and their families be arranged, but members of the community must be educated in disease and disability prevention.

Need for Services and Barriers

Undoubtedly, there is an immense need for social services among American Indians. However, there are such vast cultural differences between American Indians and non-Indians that traditional service approaches often do not meet the needs of the American Indian individuals. Culturally sensitive approaches must be employed in the service delivery to these groups. This need, however, is further complicated by the fact that there is much diversity between American Indian tribes themselves. There are 339 federally and state recognized American Indian tribes and 227 federally recognized native entities in Alaska. Furthermore, there are 175-200 languages spoken by American Indians and Alaska natives (Marshall & Largo Jr., 1999). In addition to cultural differences, socioeconomic conditions and geographic location may also hinder services (Xuequin, Coyle, Wares, & Cornell, 1999).

Many Native American communities are located on reservations, which are typically isolated from non-Indian communities and lack public transportation (Xuequin et al., 1999). This may make it difficult to access health care, as well as psychological and rehabilitation services. With few employment opportunities on or near reservations, the unemployment rate among non-disabled Indians may be quite high. The unemployment rate among the 140,000 Navajos living in rural New Mexico, Utah, and Arizona has recently been as high as 75 percent (Xuequin et al., 1999). Thus, persons may not be able to afford the services that they need, or receive health benefits that will assist in the cost of such services. The lack of employment opportunities poses difficulties for rehabilitation counseling services; establishing meaningful employment for disabled Indians is virtually impossible (Morgan, Guy, Lee, & Cellini, 1986). Lack of housing and poor roads further add to the frustrations caused by geographic barriers.

Native American Beliefs & Practices

Role of Family Family plays a significant role in service outcomes. Native American extended families are typically close knit. This extended family unit also acts as an economic and social unit (Lowrey, 1983). The extended family residence unit may consist of up to ten households of biological families that are usually located within "shouting" distance of each other (Kluckhohn & Leighton, 1996). Involvement of the family and the community in the life of the person with a disability is thus fundamental. Each individual is viewed as having a specific role to play within his or her community and/or family unit (Rocek, 1995). In fact, close-knit relationships among family members and community members are emphasized so strongly that it is sometimes believed that an individual who has few close interpersonal relationships with others is suffering from a disability of sorts (Pichette, Garrett, Kosciulek, & Rosenthal, 1999).

The importance of the family system and the impact that disability may have on the family is captured by the beliefs of the Canadian Aboriginals. For this group, an illness or disablement which affects a man's ability to work and support his family is viewed as a calamity. It can have a major impact on a man's self image: it may remove him from his traditional role and lower his status within the family and the community at large (National Aboriginal Network on Disability, 1992). The role or function of women in such families is also important. Not unlike Western non-Indian cultures, social pressures inform women that if their physical appearance is abnormal, they are flawed. Women with disabilities in this particular culture are not only viewed as unattractive, they are commonly considered to be unfit sexual partners and mothers. Aborigines prescribe the role of the woman to be the giver and nurturer of life, placing emphasis on her child bearing and child rearing roles. A disability that affects her ability to marry and have children can result in lower status among the community (National Aboriginal Network on Disability, 1992).

Religion & Spirituality The spiritual beliefs and practices of American Indian culture differ greatly from those of the more dominant American culture. Many tribes have superstitions and traditions that were handed down from generation to generation, often in the form of oral chants and stories (Lowrey, 1983). These legends occupy an important place in every-day life. There are many legendary beings in Native American beliefs. Some of the Navajo legends, for example, include the Coyote, who is regarded as one of the Holy People. Others are the Spider Woman, who taught the women to weave; the Hero Twins, who killed monsters to make the world safe; and the chinde, or malevolent ghosts of dead people inhabiting the earth (Guy, 1988). Many Navajos believe that these ghosts can cause sickness and accidents. It is also believed that witches are real people who practice black magic to bring harm to others as a form of revenge or personal gain. Achieving balance with nature is also a strong goal for most Native American tribes; falling out of balance with nature can result in disease or disability. For the Navajo, disability is believed to be rooted in the past (Martin, 1981). Parents or grandparents may be the violators of one or more natural laws, and the disability is the result. Hence, religion and moral belief is deeply embedded in Native American life, so much so that immoral behavior and disability and disease are closely linked (Guy, 1988).

The religious beliefs of Native Americans are private and very personal. According to Locust (1985), American Indians are often reluctant to discuss personal matters, particularly with non-Indians. In certain tribes, doing so may even violate tribal customs. As a result of this, little is known about Native American beliefs regarding the nature or causes of mental retardation, in spite of a prevalence rate of mental retardation almost four times the national average (Locust, 1985). In some American Indian cultures, disability is seen as a manifestation of the supernatural, the result of an evil spell, or an unnamed wrong. For some American Indian tribes, the term "disability" refers to flawed interpersonal relationships or moral wrongdoing, as opposed to physical or medical manifestations (Pichette, Garrett, Kosciulek, & Rosenthal, 1999). The most commonly regarded causes of sickness are bad dreams, excessive gambling, excessive sexual activity, ignorance of proper ceremonies, contact with the dead or death, and sorcery or witchcraft (Trennert, 1998). Any contact with a corpse or death, the Navajo believe, is detrimental. The Navajo refer to themselves as the Earth surface people, and they look to supernatural beings, the Holy People, for healing (Trennert, 1998). If the Earth surface people break a taboo, the Holy People can inflict illness on that person or their unborn children as punishment. Many tribes believe that speaking about a deformity or disability may give it power to manifest itself or express itself in human form. Native Americans are reluctant to discuss disease and disability for this reason.

In the American Indian belief system, health is not only a physical state, it is a spiritual state as well (Martin, Frank, Minkler, & Johnson, 1988). Thus, physical problems are often treated as spiritual ailments. It has been shown that more than one quarter of Native American households have at least one family-member who is functionally disabled, and more than half of all households have two (Locust, 1985). For assistance, families are often more likely to look first to natural healing methods, including medicine men, sweat lodges, sandstone ceremonies, chants, and herbal remedies, before exploring more formal social and health-related services. In fact, in some cases, healing ceremonies are believed to be the only appropriate intervention for certain medical conditions (Trennert, 1998). For those who do pursue treatment by non-Indian medical doctors, they are encouraged to also participate in natural healing interventions, to remove spells, make restitution for perceived wrongdoing, or to restore harmony and balance with oneself and their environment. These practices are further illustrated by the fact that in many Native American languages, there is no word for disability (Marshall & Largo, 1999), germ, or disease (Locust, 1985).

Other beliefs

There are other Native American beliefs and practices that merit mention with respect to the delivery of human services. First, the concept of time and punctuality is different in Native American culture. For American Indians, the reference of time is in the form of naturally occurring events, such as the rising and setting of the sun, the phases of the moon, and the seasons. Punctuality is hence not of great importance. Relationships with family and friends are more important than arriving promptly to an appointment; talking to one's relative takes precedence over making and keeping the appointment (Lewis & Ho, 1975). Second, the strong emphasis on family is also slightly different from Caucasian communities, where emphasis on individuality is stronger. Indians view their family and their community as the center of their existence, rather than their occupation, career pursuits or material possessions. Third, employment is conceptualized differently than in Caucasian culture. Occupations are passed on from generation to generation in the Indian culture. Whether the job is weaving, silversmithing, or tending sheep, work symbolizes a way of life; it is directly related to the individual's historical, cultural, and religious roots. The reasons Indians choose a job may thus be very different from why non-Indians may choose a job. These differences may undoubtedly affect a Native American's interest in receiving social services and his or her interaction with service providers.

Recommendations for Service Providers

The many aspects of Native American culture may seem to pose barriers to the delivery of human services by persons not of the same culture. However, these barriers are by no means insurmountable, and by paying close consideration to the following recommendations, service providers are likely to find that their efforts are successful. Because transportation poses a problem for many American Indians living on isolated reservations, some social services may have to be provided from within the home. Not only would doing so resolve the transportation dilemma, but it would also enable family members to be more easily included in services. Because the family is of such immense importance in Indian culture, including the family in services is likely to be very beneficial, particularly in situations where disability is construed as the result of poor interpersonal relationships.

The requirement of special training for service providers working in regions heavily populated by Native American groups might serve to increase cultural sensitivity. Familiarity with differences in language and communication, as well as religious and spiritual practices, may prepare service providers for the cultural differences that they might encounter, as well as reduce the likelihood of miscommunications. For example, understanding the Native American view of time may help explain missed appointments or appointments for which a consumer arrives late; in the absence of such diversity training, these actions might be inaccurately interpreted as laziness or irresponsibility on part of the consumer. In addition, looking directly into a person's eyes is interpreted by the Navajo as a sign of disrespect or hostility. Without such knowledge, the downward cast of a consumer's eyes might also be misinterpreted by a service provider (Trennert, 1985). Navajo recipients of human services have expressed appreciation for working with service providers who are fluent in both English and Navajo, and who are familiar with the Navajo traditions, culture, and environment (Guy, 1988). Attracting Native Americans to positions as human service providers might also enhance service delivery for Native Americans, as they may be more open to discussing private family matters with someone of similar ethnic background. Additionally, non-Indian service providers might seek an Indian "liaison" for a particular region who can facilitate their crossing of the cultural barriers. Because disability and disease are viewed as more than physical ailments, service providers must recognize, respect, and even encourage the role of spiritual ceremonies and practices. Natural approaches to healing should be incorporated into the service plan, and the service plan should thus speak to whole person, as opposed to just the physical or organic aspect of the person. Non-Indian service providers might even participate with the consumer in such healing practices and rituals.

We hope that service providers will benefit from the information included in this report, and that it will serve to increase awareness for the views of disability among Native American tribes and will inspire more culturally sensitive practices. With the prevalence of disability so high among Native Americans, it is crucial that social service personnel are knowledgeable and supportive of their ways and practices, or, sadly, the prevalence rate may continue to increase.

REFERENCES

Guy, E. (1988). Navajos' commitment benefits individuals with disabilities. *Journal of Applied Rehabilitation Counseling, 19* (4) 26-28.

Kaltenbach, C. (1980). *Health problems of the Navajo area and suggested intervention.* Window Rock, Arizona: Navajo Health Authority.

Kluckhohn, C. & Leighton, D. (1946). *The Navajo.*_New York: Doubleday & Company, Inc.,

Lewis, R. G., & Ho, M.K. *(1975).* Social work with Native Americans. *Social Work, 20 (5),* 379-382.

Locust, C.S. (1985). *American Indian beliefs concerning health and unwellness.* Tucson, AZ: University of Arizona, Native American Research and Training Center.

Lowrey, L. (1983). Bridging a Culture in Counseling. *Journal of Applied Rehabilitation Counseling 14(1),* 69-73.

Marshall, C.A., & Largo Jr., H.R. (1999). Disability and rehabilitation: a context for understanding the American Indian experience. *Lancet, 354,* 758.

Martin, *M.* (1981). Native American medicine: Thoughts for post traditional healers. *Journal of the American Medical Association, 245* (2), 14 1-143.

Martin, *W.* E., Frank. L. W., Minkler, S., & Johnson, M. (1988). A Survey of Vocational Rehabilitation Counselors Who Work with American Indians. *Journal of Applied Rehabilitation Counseling, 19(4)* 29-34.

Morgan, C. 0. Guy, E., Lee, B., & Cellini, H. R. (1986). Rehabilitation Services for American Indians: The Navajo Experience. *Journal of Rehabilitation, 52(2),* 25-31.

National Aboriginal Network on Disability. (1992). *Self-Image and Sexuality Online.* http://www.schoolnet.ca/aboriginalldisable6/self-e.html

Pichette, E. F., Garrett, M. T., Kosciulek, J. F., & Rosenthal, D. (1999). Cultural Identification of American Indians and It's Impact on Rehabilitation Services. *Journal of Rehabilitation*, 35, 3-9.

Rocek, T. R., (1995). *Navajo Multi-Household Social Units: Archaeology on Black Mesa, Arizona.* The University of Arizona Press.

Trennert, R., (1998) *White Man's Medicine: Government doctors and the Navajo, 1863-1955.* University of New Mexico Press, Albuquerque.

US Census Bureau (1980). 1980 census of the population, characteristics of the population (Volume 1, Series PC 80-1B). Washington, DC: US Department of Commerce.

Xuequin Ma, G., Coyle, C., Wares, D., & Cornell, D. (1999). Assessment of Services to American Indians with Disabilities. *Journal of Rehabilitation*, 48, 11-16.

On Differentiating Major Depression from Chronic Sadness: A Commentary

Salvatore Cullari
Lebanon Valley College

The topic of mood disorders has been of interest to healers and the general public for thousands of years. Although we have recently begun to understand more fully their etiology and have made significant strides in developing effective treatments, they remain one of the most common and debilitating psychological disorders in the world (Judd, Paulus, Wells, & Rappaport, 1996). For example, studies show that approximately 20% of the population of the United States will experience at least one significant episode of a mood disorder at some point in their lifetime (Blazer, Kessler, McGonagle, & Swartz, 1994). Furthermore, the prevalence of mood disorders seems to be increasing and the World Health Organization (WHO, 2001) estimates that by 2020, depression may be one of the leading causes of death and disability in Western cultures.

To put the WHO report into perspective, in the year 2000, about one million people world-wide committed suicide. On these, about 60% had some type of mood disorder. In the United States, suicide is the eighth most common cause of death among all age groups, and second among persons between the ages of 16 and 34. For persons with a diagnosis of major depression, about 8% will eventually commit suicide. Furthermore, according to the WHO report, by the year 2020, depressive disorders will be the leading cause of the inability to lead a productive life-style (e.g., inability to work, live independently or die prematurely).

Although we obviously have a long way to go, over the past twenty years we have made significant strides in treating depressive disorders. Part of this progress is based on effective diagnostic and classification systems. However, over the years there has been considerable debate over whether major depression is a heterogeneous (categorical) versus a homogeneous (dimensional) disorder, with supporting evidence on both sides (Feinberg, 1992; Haslam & Beck, 1994; Ruscio & Ruscio, 2000; Santor & Coyne, 2001). This distinction basically looks at whether major depression is made up of distinctive sub-categories or is really only one disorder where symptoms may vary by both type and degree on a continuum.

Author info: Correspondence should be sent to: Salvatore Cullari, Ph. D., Lebanon Valley College, Annville, PA 17003, or Cullari@lvc.edu.

North American Journal of Psychology, 2002, Vol. 4, No. 2, 317-320.
© NAJP

In terms of the heterogeneous model, the most validated distinction is that of an endogenous or biological type of depression, versus a nonendogenous or reactive type. The first is said to be due to some type of inherited predisposition and/or biological abnormality and responds best to somatic treatment such as medications or ECT (Feinberg, 1992). The second is said to be due to life events and stressors, and responds well to both somatic and non-somatic types of treatment. Somewhere in between is the diathesis-stress model, which suggests that serious mental disorders are due to both genetic predispositions (diathesis) and environmental stressors, which precipitate their onset (Monroe & Simons, 1991).

Despite empirical and clinical support for the endogenous model (see Feinberg, 1992), it tends to be inconsistent with the Diagnostic and Statistical Manual (DSM-IV-TR, American Psychiatric Association, 2001), which is the most commonly used diagnostic system for mood disorders in the US, and which focuses largely on symptoms rather than etiology. As an analogy, consider two persons entering a hospital emergency room with common symptoms of nausea, lightheadedness, racing heart and palpitations. Although the symptoms may be similar, one person may be having a heart attack while the other is having a panic attack. I doubt that anyone would recommend treating the symptoms in a similar manner without considering the etiology. There may be a parallel in the diagnosis and treatment of depressive disorders.

The purpose of this paper is to promote an etiological based model of diagnosing depressive disorders and to introduce the contrast of major depression (MD) and chronic sadness (CS) as an example of how this system might be used. The description of these two possible discrete entities is based on the clinical experience of the author, but it is hoped that this paper will spur more empirically based research, and ultimately to more effective treatments for various depressive disorders as a whole.

Table 1 summarizes the similarities and possible differences between these two conditions that have been subjectively noted by the author. MD and CS share many of the same symptoms and severity levels except for psychosis. This observation is consistent with the endogenous-nonendogenous view of depression, but is it not totally contradictory to the diathesis-stress model. At the same time, CS is not the same as the relatively mild and brief reactive depression that everyone experiences at times, or an adjustment disorder, which by definition can only last six months (American Psychiatric Association, 2001).

In CS, there is always a precipitating event, which includes some type of significant (real or perceived) loss. This may include the death of a loved one, divorce, loss of bodily function, e.g., amputation, serious brain injury or a serious physical disorder, or a considerable life-style change. In many ways, CS resembles a grief reaction except that it may be precipitated by other losses and at times may persist even longer.

Typically, the premorbid functioning of persons with CS is normal, and there is usually no history of severe depression prior to the precipitating event. Although CS may be a chronic condition, it is different than a dysthymic disorder in that onset occurs only after a clear precipitator, and relapse is rare after resolution. In terms of treatment, it tends to respond more readily to psychological interventions rather than medications and in fact, medications tend to be ineffective with CS.

It is hoped that this preliminary differential diagnostic scheme will lead to an open dialogue, more objective and controlled studies and ultimately to more effective treatments. For example, I would recommend that research investigating treatment interventions for major depression divide subject groups according to etiology in order to help determine differential effects. In addition, future studies might also look at the results of depression rating scales based on the schema presented above.

REFERENCES

American Psychiatric Association. (2001). *Diagnostic and statistical manual of mental disorders.* (4th ed., Text Revision). Washington, DC: Author.

Blazer, D. G., Kessler, R. C., McGonagle, K. A., & Swartz, M. S. (1994). The prevalence and distribution of major depression in a national community sample: The National Comorbidity Survey. *American Journal of Psychiatry, 151*, 979-986.

Feinberg, M. (1992). Subtypes of depression and response to treatment. *Journal of Consulting and Clinical Psychology, 60*, 670-674.

Haslam, N., & Beck, A.T. (1994). Subtyping major depression: A taxometric analysis. *Journal of Abnormal Psychology, 103*, 686-692.

Judd, L. L., Paulus, M. P., Wells, K. B., & Rapaport, M. H. (1996). Socio-economic burden of subsyndromal depressive symptoms and major depression in a sample of the general population. *American Journal of Psychiatry, 153*, 1411-1417.

Monroe, S.M., & Simons, A.D. (1991). Diathesis—Stress theories in the context of life stress research: Implications for the depressive disorders. *Psychological Bulletin, 110*, 406-425.

Ruscio, J., & Ruscio, A. M. (2000). Informing the continuity controversy: A taxometric analysis of depression. *Journal of Abnormal Psychology, 109*, 473-487.

Santor, D. A,, & Coyne, J. C. (2001). Examining symptom expression as a function of symptom severity item performance on the *Hamilton Rating Scale for Depression. Psychological Assessment, 13*, 127-139.

WHO (2001). *The world health report 2001: Mental health: New understanding, new hope*. Geneva: World Health Organization.

TABLE 1 Symptom Differences between Major Depression (MD) and Chronic Sadness (CS)

SYMPTOMS	MD	CS	INCREASED SEVERITY OF SYMPTOMS
Depressed mood	yes	yes	No difference
Suicidal thoughts/gestures	yes	yes	No difference
Anhedonia - appetite disturbance	yes	yes	MD
Loss of libido	yes	yes	MD
Appetite disturbance- weight loss	yes	yes	MD
Fatigue	yes	yes	MD
Sleep disturbance	yes	yes	CS *
Hopelessness	yes	yes	CS
Helplessness	yes	yes	CS
long duration	yes	yes	CS
Psychosis	yes	no	
Discrete episodes	yes	no	
Cognitive impairment	yes	no	
Gender differences	yes	no	
Morbid obsession with death	yes	no	
Psychomotor disturbance	yes	no	
Reduced self-esteem	yes	no	
Inability to function	yes	no	
Obvious precipitator	sometimes	always	
Relapse after recovery	likely	unlikely	

*typically involves middle of the night awakening rather than early morning

BEST COPY AVAILABLE

How to Submit Papers to *NA*JP

The submission of good papers is very important to any psychology journal. We hope that you will support our views about what a psychology journal should be like by submitting papers now. We will try to mail a decision on your paper within 5 weeks of receipt by us. We will publish your paper in an average of 5 months of receipt of a final accepted version.

Submit 1 paper copy of each manuscript prepared according to APA guidelines. Also include a diskette with your manuscript or (preferably) e-mail the file to NAJP@mpinet.net. Include 2 large stamped envelopes so we can send your manuscript out for review. We encourage you to have your paper read by peers before submission.

If the final paper is accepted, please send a check or purchase order for the amount specified with your final draft (billed at $25 dollars per typeset page in units of two pages).

Submission of a manuscript to the Journal is a contract to publish with the Journal if the paper is accepted. Because of our rapid editorial feedback, simultaneous submission to other journals is not allowed. We will offer brief comments on the suitability of a topic for publication in *NA*JP before a submission.

Send all correspondence to: *North American* Journal of Psychology, 240 Harbor Dr., Winter Garden, FL 34787, phone 407-877-8364.

SUBSCRIPTIONS

All subscriptions are on a calendar year basis only. If you order late in the year, the earlier issues will be sent to you. Institutional rate, $65. Individual rates, paid by check, $30. Add $7 for Canadian subscriptions. Add $8 for subscriptions outside North America. Phone orders accepted with Visa, MasterCard or American Express. Call 407-877-8364. The *North American* Journal of Psychology is now published semi-annually in June & December.

Visa ☐ AMEX ☐ MC ☐

Card # Exp. Date

Name

Address

City State Zip

A Summary of Our Philosophy

The ***North American* Journal of Psychology** publishes scientific papers that are of general interest to psychologists and other social scientists. Papers may encompass a wide variety of topics, methods, and strategies. It is our belief that many good scientific papers are rejected for invalid reasons. Some journals reject up to 90% of their submissions because they don't have space. Furthermore, studies of the review process strongly suggest that not all of the lucky 10% are better papers than those that get rejected. In our view many of those rejected papers are meritorious with need for only minor revisions. Perhaps our philosophy can be summed up this way: we are looking for ways to strengthen good papers rather than searching for weak excuses to reject them.

North American

Journal of Psychology

'ol. 4, No. 3 2002

Advisory Editorial Board

ADRIAN FURNHAM
Psychology
University College London

MANUEL LONDON
Psychology
SUNY, Stony Brook

DOUGLAS JACKSON
Psychology
University of Western Ontario

KAREN LYNESS
Management
Baruch College, CUNY

JAMES M. JONES
Psychology
University of Delaware

ROBERT PASNAK
Psychology
George Mason University

HEIDI KELLER
Psychology
University of Osnabrueck

RICHARD RYAN
Psychology
University of Rochester

PHILIP ZIMBARDO
Psychology
Stanford University

Editor

LYNN E. McCUTCHEON

How to Subscribe to The *North American* Journal of Psychology

All subscriptions are on a calendar year basis only. If you order late in the year, the earlier issue will be sent to you. Institutional rate: $65. Individual rate: $30. Add $7 for Canadian subscriptions. Add $8 for subscriptions outside North America. *North American* Journal of Psychology is published in March, June and December.

Now visit our website at http://najp.8m.com

North American Journal of Psychology

December, 2002 ISSN 1527-7143 Vol. 4, No.3

Girls with ADHD and Associated Behavioral Problems: Patterns of Comorbidity 321
Amy S. Kerivan Marks, Mary Nichols, Christine Blasey, Pamela Kato, & Lynne Huffman

Narcissism and Aggression: Is Inflated Self-esteem Related to Aggressive Driving? 333
George E. Schreer

Positive Attitude Towards Dreams: Reliability and Stability of a Ten-item Scale 343
Michael Schredl, Constanze Brenner, & Christina Faul

The Impact of Sport Team Identification and Attributions of Ability and Effort on Spectators' Impressions of Athletic Performance 347
Daniel L. Wann, Brian L. Keenan, Steve Burnett, Jennifer Martin, Leslie Page, Lori Smith, & Christopher D. Lantz

An Interview With Albert Ellis About Rational Emotive Behavior Therapy 355
Albert Ellis, Michael F. Shaughnessy, & Virginia Mahan

Counselor-client Matching on Ethnic, Gender, and Language: Implications for Counseling School-aged Chidren 367
Julie Hall, Dana Kaplan, Howard B. Lee, & Steven G. Little

Parent Discipline Scale: Parental Discipline Styles as a Function of Transgressor Type 381
N. L. Lopez, H. G. Schneider, & C. S. Dula

The Effect of Autobiographical Writing on the Subjective Well-being of Older Adults 395
Nancy Richeson & James A. Thorson

Psychology Training Regarding HIV/AIDS Revisited 405
Carlos Escoto

Attachment Styles, View of Self, and Negative Affect 417
Amy Van Buren & Eileen L. Cooley

Assessing Attitude Toward Christianity among Chinese Speaking Adolescents in Hong Kong: The Francis Scale 431
Leslie J. Francis, Christopher Alan Lewis, & Peter Ng

Young Females' Perceptions of the Impact of a Sexual Abuse Experience: A Focus Group Approach 441
Ruth S. Buzi, Susan R. Tortolero, Peggy B. Smith, Michael W. Ross, & Robert E. Roberts

Mood, Rumination, and Mood Awareness in the Athletic Performance of Collegiate Tennis Players 457
Vann B. Scott, Kamee B. Stiles, David B. Raines, & Andreas W. Koth

A Lottery in the Bible Belt: Factors Affecting Vote Decisions 469
Larry Powell & William R. Self

An Interview with Sonia Nieto about Multiculturalism 479
Bill Gaedke & Michael F. Shaughnessy

Straight From the Heart: Integrating Service Values with the Study of Psychology 489
Judith Reifsteck

Editor's Comments

Thanks again to all the reviewers, all of whom did a fine job of critiquing papers promptly and with great insight. Without a doubt our reviewers have made this journal what it is today. In the previous issue, I recognized three reviewers by name who have done a fine job of reviewing two or more papers for NAJP. For this issue a tip of the editor's cap goes to three more reviewers, specifically Elizabeth Brazelton, Steve Bridges, and Felicity Allen. Elizabeth was an old friend of mine from our days together in graduate school at Auburn. She passed away several months ago. Elizabeth will be missed by all who knew her.

I am sure readers will enjoy the candid remarks made by Albert Ellis in the interview done on behalf of NAJP by Michael Shaughnessy. There are also fascinating articles on girls who have ADHD (see Marks, et al), attitudes toward dreaming (see Schedl, et al), and the relationship between narcissm and aggressive driving (see Schreer). And these are just the first three articles.

In the last issue and the one before I applauded the willingness of certain APA editors to make their journals more "friendly." I want to continue to encourage other editors to do the same.

Lynn E. McCutcheon, editor

Girls With ADHD And Associated Behavioral Problems: Patterns of Comorbidity

Amy S. K. Marks
Children's Health Council, Stanford University

Mary Nichols
Children's Health Council, Palo Alto, CA

Christine Blasey Pamela M. Kato Lynne C. Huffman
Stanford University

Attention Deficit Hyperactivity Disorder (ADHD) is a behavioral disorder commonly seen by pediatricians, psychiatrists and community mental health specialists, and is less frequently identified in girls. This review investigated parent-reported behavioral profiles of girls among different ADHD diagnostic comorbid groups. Demographic and *Child Behavior Checklist* (CBCL) data were available for 40 girls and a comparison group of 55 boys. Independent measures *t*-tests compared CBCL *T*-scores among three diagnostic groups: (1) girls with ADHD only; (2) girls with ADHD and co-occurring learning disabilities (LD); and, (3) girls with ADHD and other co-occurring DSM Axis I diagnoses. Girls in the ADHD/Axis I group had the most clinically significant behavioral problems; in contrast, girls in the ADHD/LD group had few behavioral symptoms. This pattern was not found for the boys' groups, suggesting that the co-occurrence of ADHD and other behavioral problems, especially LD, may have a different behavioral phenotype in girls than in boys.

Attention Deficit Hyperactivity Disorder (ADHD) is a common childhood disorder, with epidemiological studies revealing prevalence rates ranging from 4% to 12% in the general population of 6 to 12 year olds (Brown, 2000; Brown et al., 2001). Despite the overall frequency with which children having ADHD are seen by pediatric mental health and primary care practitioners, it is a behavioral condition that is underidentified in girls (Biederman, 1998; Biederman et al., 1999). Research addressing the behavioral manifestations of ADHD in females has been limited (Arnold, 1996). Further characterization of the disorder in girls is critical in order to optimize our diagnostic and treatment strategies.

Author info: Correspondence should be sent to: Amy Marks, Children's Health Council, Dept. of Outcomes Measurement and Research, 650 Clark Way, Palo Alto, CA 94304. Phone 650-617-3850. Email akerivan@stanford.edu

North American Journal of Psychology, 2002, Vol. 4, No. 3, 321-332.

© NAJP

In at least fifty percent of ADHD diagnoses, children also suffer at least one other psychiatric disorder, and an estimated 20-25% of children with ADHD receive a co-diagnosis of learning disorder (LD; Brown, 2000). Our understanding of these co-occurring disorder processes is limited by poor measurement strategies, clinical disciplinary biases, and the unpredictable sequence of emerging behavioral symptoms (Cantwell & Baker, 1991; Shaywitz, Fletcher, & Shaywitz, 1995).

Further, while it is established that other behavioral, emotional, and learning problems significantly co-occur with ADHD, (Purvis & Tannock, 1997; Shaywitz et al., 1995) the concurrent behavioral manifestations of ADHD and other disorders have been described in predominantly male samples. For example, several studies have demonstrated that boys with ADHD and comorbid LD typically demonstrate externalizing behavior problems, as well as language-related difficulties, (Beitchman & Young, 1997; Brown, 2000; Faraone et al., 1993; Hinshaw, 1992) and peer relation problems (Flicek, 1992). In addition, research has shown that boys with ADHD and co-occurring LD often exhibit similar, if not higher, levels of parent-reported externalizing problem behaviors than boys with ADHD alone (Nigg, Carte, Hinshaw, & Treuting, 1998; Piesco, Baker, Silva, & Brooke, 1996). In contrast, much less is known about the behavioral symptoms of girls with comorbid ADHD (Newcorn et al., 2001; Rucklidge & Tannock, 2001; Sharp et al., 1999 as exceptions). One recent study suggests that girls and boys with ADHD are very similar in presentation (e.g., Sharp et al., 1999) while others have found that girls with ADHD are less symptomatic than boys, particularly with regard to activity level and aggression (e.g., Carlson, Tamm, & Gaub, 1997; Newcorn et al., 2001). Given these conflicting findings, a lack of attention to LD as a comorbid condition of ADHD, and the relative dearth of information regarding ADHD in females, additional research that addresses girls with ADHD and co-occurring behavioral and learning problems clearly is needed.

In attempts to more thoroughly understand comorbidity issues in children with ADHD, it has been suggested that the *Child Behavior Checklist* (CBCL; Achenbach, 1991a) might be useful for behaviorally characterizing and discriminating comorbid and non-comorbid cases of ADHD (Biederman et al., 1993). The use of behavior rating scales such as the CBCL to further characterize the Diagnostic and Statistical Manual of Mental Disorder's (DSM; *Diagnostic and Statistical Manual of Mental Disorders*, 1994) ADHD diagnoses is well supported (Biederman et al., 1993; Eiraldi, Power, & Karustis, 2000). The American Academy of Pediatrics recently presented clinical practice guidelines for ADHD, recommending that, although behavioral symptoms are only one of the six diagnostic criterion for an ADHD diagnosis, tests such as the CBCL may be useful in differentiating children with ADHD from children with other behavioral problems (Committee on Quality Improvement, 2000). Again, however, the majority of studies evaluating relations between the CBCL scales and DSM diagnoses have focused on male pediatric samples and have not attempted to discriminate and identify behavioral characteristics among comorbid ADHD cases in females.

The purpose of the current study is to further investigate the behavioral symptomatology of girls with ADHD, girls with co-occurring ADHD and LD, and co-occurring ADHD and other Axis I disorders. Informed by conclusions from previous research on boys with ADHD, we hypothesized that girls with ADHD and Axis I disorders and girls with ADHD and LD would have higher levels of parent-rated problem behaviors than girls with ADHD alone. To fully explore these hypotheses, a comparison group of boys with similar "pure" and comorbid diagnoses also was examined.

METHOD

Participants

The target sample was comprised of school-aged children and adolescents who were evaluated for attention problems in a non-profit community behavioral health agency located in Northern California. All children who had medical charts and were referred for an evaluation of ADHD between 1988 and 1998 ($N = 179$) were included. In one case, a diagnosis of ADHD made by a community clinician more than one year prior to the participant's presentation to our agency precluded participation (because the basis and method of diagnosis were unknown to this agency). As part of the agency's client intake process, demographic information and CBCLs were requested of all parents of new clients presenting to the agency's ADHD Clinic for an evaluation of ADHD. All participants receiving the full diagnosis of ADHD (n=133) and who had complete demographic, medical history, and CBCL data reported by parents (n=95) were included and recorded for a female group (n= 40) and an age and ethnicity-comparable comparison group of boys (n=55). Age, ethnicity, and referral sources for the final study sample are presented in Table 1.

Our sample consisted of predominantly Caucasian and Hispanic preadolescent children referred by a pediatrician or family member. Although many participants chose not to report family income information, the location of residence (by zip code and street address) indicated that over 90% of participants were from middle- and upper-class communities.

TABLE 1. Demographic Characteristics of the Sample

	Girls group	Boys comparison group
N	40	55
Mean age in years (SD)	9.0 (3.2)	8.9 (3.2)
Ethnicity:		
Caucasian	n=20 (50.0%)	n=25 (45.5%)
Latino/a	n=4 (10.0%)	n=3 (5.5%)
Other	n=3 (7.5%)	n=1 (1.8%)
Missing	n=13 (32.5%)	n=26 (47.3%)
Referral sources		
Pediatrician	n=13 (32.5%)	n=16 (29.9%)
Family member or parent	n=13 (32.4%)	n=14 (25.5%)
Therapist or counselor	n=7 (16%)	n=10 (18.2%)
School or educator	n=4 (9%)	n=7 (12.7%)
Lawyer/courts	n=1 (2.3%)	-
Not available	n=2 (4.5%)	n=8 (14.5%)

Measures

A trained research assistant collected data from the following measures during the chart review:

Diagnostic and Statistical Manual of Mental Disorders (DSM III-R Diagnostic and Statistical Manual of Mental Disorders, 1987; and IV Diagnostic and Statistical Manual of Mental Disorders, 1994). All diagnoses were made by one of three licensed psychologists staffing the ADHD Clinic, and were based upon a semi-structured clinical interview using the DSM IV (75%) or III-R (25%) criteria. Participants were given the full diagnosis of ADHD and any co-occurring disorders only if all criteria from the DSM were met unequivocally. Likewise, Learning Disorders also were made according to the DSM, and included Reading Disorder, Mathematics Disorder, and Disorder of Written Expression; these disorders were only given if all DSM criteria were met.

Child Behavior Checklist. (Achenbach, 1991a). The CBCL is a widely used and well-standardized measure of child behavioral symptomatology. This parent-reported 113-item checklist is used to produce standardized profiles of behavior that may assist in the identification and assessment of a wide range of child behavioral disorders. The checklist is age- and gender-standardized and is designed to describe the behavior of children aged 4-18 years. The CBCL has substantial test-retest reliability, inter-rater agreement, internal consistency and external validity (Achenbach, 1991b). For the purposes of this chart review, we recorded *T*-scores for three total problems scales (i.e., Total Problems, Internalizing Problems and Externalizing Problems) as well as eight CBCL syndrome scales (i.e., Attention Problems, Somatic Complaints, Withdrawn, Anxiety-Depression, Social Problems, Thought Problems, Delinquent Behavior, and Aggressive Behavior). For the syndrome scales, a T-score greater than or equal to 67 is clinically significant, representing either 'borderline clinical' or 'clinical' levels of symptoms. Mothers represented 83% of respondents for the girls and 87% of respondents for the boys. Other respondents included fathers, adoptive parents, foster parents and grandparents. Eight percent of CBCLs were filled out by both mothers and fathers; in these cases, the mothers' reports were used.

Procedure

A clinical chart review was conducted to acquire the CBCL data, diagnoses, and demographic information. *Diagnostic and Statistical Manual* multi-axial diagnoses were obtained from formal mental health evaluations, and were utilized to create three diagnostic groups: (1) girls diagnosed only with ADHD ("ADHD" group); (2) girls diagnosed with ADHD and learning disabilities ("ADHD/LD" group); and, (3) girls diagnosed with ADHD and other DSM Axis I diagnoses (e.g., adjustment disorder, anxiety disorder, oppositional defiant disorder, and dysthymia; "ADHD/Axis I" group). One girl had diagnoses of ADHD, LD, and another Axis I disorder; she was included in the ADHD/Axis I group due to the strong influence of her other disorder on her behavioral symptomatology. Diagnostic groups similarly were constructed for a comparison group of boys. There were no cases where LD and Axis I disorders were coexisting in the boys' group. CBCLs are considered stable for an individual over the period of at least 1 year (Achenbach, 1991b), and had been administered in the ADHD Clinic within six months of all of the participants' ADHD diagnosis/evaluation.

Statistical Analyses

In order to establish normal distributions, CBCL scores were transformed logarithmically. Thereafter, independent samples t-tests were used to compare the transformed CBCL T-scores across the three female groups: ADHD, ADHD/LD, and ADHD/Axis I. Similar analytic strategies were applied to the comparison male group CBCL data. An alpha of .01 (two-tailed) was used as the threshold for statistical significance.

RESULTS

Gender comparisons

Girls and boys were similar in age ($t_{(93)} = .01, p=.99$), with a mean age of 8.9 (*SD* 3.1) for girls and 8.9 (*SD* 3.2) for boys. The frequency distributions of the three diagnostic categories were not statistically different for girls and boys (two-way $X^2=1.1$, $p=.57$). Mean scores for CBCL syndrome scales were similar for girls and boys, with the exception of the Attention Problems syndrome scale (mean Attention Problems *T*-score for girls=72 and for boys=67; $t_{(93)} = 2.5, p=.014$).

Diagnostic Group Distributions

Diagnoses of Attention Deficit Hyperactivity Disorder included 25% DSM IIIR-based and 75% DSM IV-based Axis I ADHD[1]. For the study group of 40 girls, there were eight (20%) in the ADHD group, 14 (35%) in the comorbid ADHD/LD group, and 18 (45%) in the comorbid ADHD/Axis I group. For the comparison group of 55 boys, there were 11 (22%) in the ADHD group, 14 (28%) in the comorbid ADHD/LD group, and 30 (60%) in the comorbid ADHD/Axis I group.

Behavioral Symptomatology: CBCL T-score Results by Diagnostic Group

For the three diagnostic groups, mean *T*-scores on the CBCL Attention Problems syndrome scale were in the clinical range for all three diagnostic groups: ADHD (70.0), ADHD/LD (72.8), and ADHD/Axis 1 (71.9). On the Aggressive Behavior syndrome scale, the mean scores for the ADHD (67.5) and the ADHD/Axis I (71.6) groups were in the clinical range. Lastly, the mean *T*-scores for the Social Problems syndrome scale were in the clinical range for the ADHD diagnostic group (68.0). All other mean syndrome scale scores for the girls' diagnostic groups were in the non-clinical range, and are available upon request from the first author.

For the boys' groups, CBCL mean *T*-scores on the Attention Problems syndrome scale reached the clinical range for the ADHD (69.3) and ADHD/LD (71.3) groups only. All other mean CBCL syndrome *T*-scores for the boys' diagnostic groups were in the non-clinical range.

---

[1] In the DSM IV, the diagnosis of ADHD includes subtypes of Inattentive Type, Hyperactive-Impulsive Type, and Combined Type.

Between - Diagnostic Group CBCL Paired Comparisons for Girls and Boys

For the girls, independent samples *t*-tests indicated that the ADHD and the ADHD/LD groups differed significantly on the Somatic Complaints and approached a significant difference on the Delinquent Behavior syndrome scales ($t_{(20)} = 2.71$, $p=.013$, effect size 21% and $t_{(20)} = 2.32$, $p= .031$, effect size 10%, respectively), indicating more problems in these domains for girls with ADHD only than for girls with a comorbid diagnosis of ADHD/LD. No other ADHD and ADHD/LD group dif- ferences were observed.

Comparisons of the ADHD/LD and ADHD/Axis I diagnostic groups indicated that girls with ADHD and Axis I diagnoses were rated significantly higher on the Somatic Complaints and approached a statistical difference on the Aggressive Behavior scales only ($t_{(30)} =-2.96$, $p=.006$, effect size 21% and $t_{(30)} =-2.21$, $p=.035$, effect size 13% respectively). No other group differences were found between the ADHD/LD and ADHD/Axis I groups.

No significant group differences were found between the ADHD and ADHD/Axis I diagnostic groups on the eight CBCL syndrome scales. The *t*-test results are presented in Table 2.

The same statistical procedures were used to compare CBCL scores across the three comparison boys' diagnostic groups. The *t*-test statistics revealed no group differences on the eight CBCL syndrome scales, indicating that the comparison boys have similar parent-reported behavioral profiles, regardless of their status as having ADHD only, ADHD with comorbid LD, or ADHD with other Axis I disorders.

TABLE 2 Paired Comparisons of Girls' CBCL *t*-scores Between Diagnostic Groups.

CBCL syndrome scale	ADHD vs. ADHD/LD *t*-test statistic	ADHD/LD vs. ADHD/Axis I *t*-test statistic	ADHD/Axis I vs. ADHD *t*-test statistic
Somatic complaints	2.712*	-2.958*	-.016
Aggression	1.275	-2.208	-.726
Delinquency	2.322	-1.463	.626
Depression	.631	-1.426	-.511
Withdrawn	-1.883	.512	-1.756
Social problems	1.171	-.287	1.150
Thought problems	-.443	.269	-.243
Attention problems	-.061	.259	-.479

* two-tailed significance, p<.01

DISCUSSION

This study contributes to the growing pediatric and psychological literature on girls with ADHD and other co-occurring behavioral disorders. Using a medical chart review, we demonstrated that girls with a diagnosis of "pure" ADHD or ADHD with co-occurring Axis I disorders have several different behavioral characteristics than girls with co-occurring ADHD and LD. Across all diagnostic groups, parents rate their daughters as demonstrating clinically significant levels of attention problems. Beyond this, girls with ADHD only and ADHD/Axis I disorders were rated as having clinically elevated aggression symptoms. In addition, girls with ADHD/LD were perceived as relatively free of behavioral symptoms (with the exception of inattention).

Girls in the "pure" ADHD group were rated as having significantly higher scores on the Delinquent Behavior and Somatic Complaints syndrome scales than girls in the ADHD/LD group. Further, girls with ADHD/Axis I disorders were rated by their parents as having higher levels of aggressive behavior and somatic complaints than girls with ADHD/LD. The elevated externalizing symptoms reported by parents of girls in the ADHD and ADHD/Axis I groups are reminiscent of the ADHD children with comorbid Oppositional Defiant Disorder/Conduct Disorder reported by Newcorn et al. (2001) in their analysis of data from 498 children studied in the National Institute of Mental Health Collaborative Multisite Multimodal Treatment project. As noted by Newcorn and others (e.g., Abikoff, Courtney, Pelham, & Koplewicx, 1993), this may reflect a diagnostic bias with an increased identification of ADHD symptoms in children who are oppositional and disruptive.

Mean scores on the Somatic Complaints scale were among the lowest of the syndrome scales; nonetheless, this CBCL dimension was an interestingly apt scale for differentiating the groups with ADHD only, ADHD/LD or ADHD/Axis I disorders. One recent report has suggested that the presence of somatic complaints in girls is strongly associated with emotional disorders such as anxiety (Egger, Costello, Erkanli, & Angold, 1999). Though somatization is recognized in the DSM-IV as a criterion for a variety of anxiety disorders, there have been relatively few studies examining its utility as an important behavioral symptom in diagnosing or characterizing other mental disorders. Our findings suggest that girls with pure ADHD or ADHD/Axis I may report higher levels of somatizing symptoms, (e.g., nausea, stomach and headaches, dizziness, being tired) than girls with ADHD and LD. These findings merit further investigation and consideration, particularly in the context of primary pediatric healthcare, where somatizing symptoms may be elicited and recognized most easily.

Contrary to our hypothesis, we found that the girls in the comorbid ADHD/LD group were relatively asymptomatic. The only parent-reported behavioral symptoms of clinical significance were those on the Attention Problems scale. In this group it may be that inattention is more reasonably considered an indicator of learning problems than a manifestation of an attention deficit hyperactivity syndrome. Pursuing this formulation, a primary care pediatrician might prioritize further psychological and educational assessments.

Our findings from the boys' comparison groups correspond with most existing literature describing the behavioral symptomatology of boys with ADHD and co-occurring LD. As an exception, one previous report, based mainly on boys and using discriminant-function analyses, concluded that CBCL Aggressive Behavior and Externalizing scores could differentiate among and between pure ADHD, pure LD and comorbid ADHD/LD groups (Robins, 1992). In our study, similar results differentiating the ADHD from the ADHD/LD groups were found, but for the girls' diagnostic groups only.

Given the findings of our study, we believe that behavior ratings may be helpful in enriching our understanding of girls presenting with ADHD symptoms to community mental health care practitioners. Diagnostic group behavior patterns may have important clinical implications for ADHD identification and evaluation, may inform our understanding of ADHD etiology, and can shape decision-making about appropriate treatment strategies.

We recognize that the retrospective design of this study increases its limitations. As we could not predetermine the source of our study sample recruitment, our sample contained a combination of pediatric, mental health specialist, and self-referred participants. Our sample was not large enough to evaluate symptomatology by diagnosis and referral source, limiting our ability to draw broader conclusions. Gaub and Carlson (Gaub & Carlson, 1997) have demonstrated that clinic referred girls and boys with ADHD exhibited similar levels of inattention, internalizing behavior, and peer aggression, while non-clinic referred girls and boys with ADHD appeared notably different in their behavioral symptomatology. Concerning the validity of our parent-respondent data, recent research supports the use of parent ratings of the CBCL, in particular the *Attention Problems* and *Aggressive Behavior* scales, in identifying behavioral trends in children with ADHD, with no evidence of rater biases and rater differences by child gender (Hudziak et al., 2000). In addition, CBCL parent ratings of behavior have been shown to correspond very well with structured diagnostic interview-based diagnoses. (Ostrander et al., 1998; Steingard et al., 1992). Lastly, though our sample size was only moderate, we did have similar numbers of girls and boys represented in our analyses, and we employed a limited number of statistical tests when investigating group differences.

Future research is needed to more fully address the relationship between comorbid and non-comorbid ADHD and the corresponding behavioral symptoms in girls. This information may reveal important behavioral patterns that could be utilized in future diagnostic and screening strategies. Longitudinal studies addressing the patterns of emerging behavioral symptoms also are needed, particularly in younger girls before ADHD and LD symptoms become impairing. Our findings with the Somatic Complaints syndrome scale serve as a reminder that even sub-clinical behaviors should be considered in understanding and identifying children with varying comorbid behavioral problems. Future research addressing the somatizing aspects of behavior and ADHD in girls is needed. In addition, future studies should utilize multiple respondent study designs, incorporating parent, teacher and child self-report measures in the process of assessing behavioral phenotypes of ADHD and co-occurring learning or other mental problems. Lastly, future research in this arena utilizing larger sample sizes will help us identify more subtle behavioral differences among girls and boys with comorbid and ADHD only, and will aid us to measure and predict future clinical and academic outcomes for children presenting with ADHD symptoms.

REFERENCES

Abikoff, H. B., Courtney, M., Pelham, W. E., & Koplewicx, H. S. (1993). Teachers' ratings of disruptive behaviors: the influence of halo effects. *Journal of Abnormal Child Psychology, 21*, 519-533.

Achenbach, T. (1991a). *Child Behavior Checklist for Ages 4-18*. Burlington, VT: University of Vermont.

Achenbach, T. (1991b). *Manual for the Child Behavior Check-list/4-18 and 1991 Profile*. Burlington, VT: University of Vermont Department of Psychiatry.

Arnold, L. (1996). Sex differences in ADHD; conference summary. *Journal of Abnormal Child Psychology, 24*(5), 555-569.

Beitchman, J. H., & Young, A. R. (1997). Learning disorders with a special emphasis on reading disorders: A review of the past 10 years. *Journal of the American Academy of Child & Adolescent Psychiatry, 36*(8), 1020-1032.

Biederman, J. (1998). Attention-deficit/hyperactivity disorder: a life-span perspective. *Journal of Clinical Psychiatry, 59*(7), 4-16.

Biederman, J., Faraone, S. V., Doyle, A., Lehman, B. K., Kraus, I., Perrin, J., & Tsuang, M. T. (1993). Convergence of the Child Behavior Checklist with structured interview-based psychiatric diagnoses of ADHD children with and without comorbidity. *Journal of Child Psychology, Psychiatry & Allied Disciplines, 34*(7), 1241-1251.

Biederman, J., Faraone, S. V., Mick, E., Williamson, S., Wilens, T. E., Spencer, T. J., Weber, W., Jetton, J., Kraus, I., Pert, J., & Zallen, B. (1999). Clinical correlates of ADHD in females: findings from a large group of girls ascertained from pediatric and psychiatric referral sources. *Journal of the American Academy of Child & Adolescent Psychiatry, 38*(8), 966-975.

Brown, R. T., Freeman, W. S., Perrin, J., Stein, M., Amler, R. W., Feldman, H. M., Pierce, K., & Wolraich, M. L. (2001). Prevalence and assessment of Attention-Deficit/Hyperactivity Disorder in primary care settings. *Pediatrics, 107*(3), E43.

Brown, T. E. (2000). *Attention-deficit disorders and comorbidities in children, adolescents, and adults* (First ed.). Washington D.C.: American Psychiatric Press, Inc.

Cantwell, D. P., & Baker, L. (1991). Association between attention deficit-hyperactivity disorder and learning disorders. *Journal of Learning Disabilities, 24*(2), 88-95.

Carlson, C., Tamm, L., & Gaub, M. (1997). Gender differences in children with ADHD, ODD, and co-occurring ADHD/ODD identified in a school population *Journal of the American Academy of Child & Adolescent Psychiatry, 36*(12), 1706-1714.

Committee on Quality Improvement, S. o. A.-D. H. D. (2000). Clinical practice guideline: Diagnosis and evaluation of the child with Attention-Deficit/Hyperactivity Disorder. *Pediatrics, 105*(5), 1158-1170.

Diagnostic and Statistical Manual of Mental Disorders. (Fourth ed.)(1994). Washington, D.C.: American Psychiatric Association.

Diagnostic and Statistical Manual of Mental Disorders. (Third-Revised ed.)(1987). Washington, D.C.: American Psychiatric Association.

Egger, H. L., Costello, E. J., Erkanli, A., & Angold, A. (1999). Somatic complaints and psychopathology in chilren and adolescents: Stomach aches, musculoskeletal pains, and headaches. *Journal of the American Academy of Child & Adolescent Psychiatry, 38*(7), 852-860.

Eiraldi, R. B., Power, T. J., & Karustis, J. L. (2000). Assessing ADHD and comorbid disorders in children: The Child Behavior Checklist and the Devereux Scales of Mental Disorders. *Journal of Clinical Child Psychology, 29*(1), 3-16.

Faraone, S. V., Biederman, J., Lehman, B. K., Spencer, T. J., Norman, D., Seidman, L. J., Kraus, I., Perrin, J., Chen, W. J., & Tsuang, M. T. (1993). Intellectual performance and school failure in children with attention deficit hyperactivity disorder and their siblings. *Journal of Abnormal Psychology, 102*(4), 616-623.

Flicek, M. (1992). Social status of boys with both academic problems and attention-deficit hyperactivity disorder. *Journal of Abnormal Child Psychology, 20*(4), 353-366.

Gaub, M., & Carlson, C. (1997). Gender differences in ADHD: A meta-analysis and critical review. *Journal of the American Academy of Child & Adolescent Psychiatry, 36*(8), 1036-1045.

Hinshaw, S. P. (1992). Academic underachievement, attention deficits, and aggression: Comorbidity and implications for intervention. *Journal of Consulting and Clinical Psychology, 60*(6), 893-903.

Hudziak, J. J., Rudiger, L. P., Neale, M. C., Heath, A. C., & Todd, R. D. (2000). A twin study of inattentive, aggressive, and anxious/depressed behaviors.

Journal of the American Academy of Child & Adolescent Psychiatry, 39(4), 469-476.

Newcorn, J. H., Halperin, J. M., Jensen, P. S., Abikoff, H. B., Arnold, E., Cantwell, D., Conners, C. K., Elliott, G. R., Epstein, J. N., Greenhill, L. L., Hechtman, L., Hinshaw, S. P., Hoza, B., Kraemer, H. C., Pelham, W. E., Severe, J. B., Swanson, J. M., Wells, K. C., Wigal, T., & Vitiello, B. (2001). Symptom profiles in children with ADHD: Effects of comorbidity and gender. *Journal of the American Academy of Child & Adolescent Psychiatry, 40*(2), 137-146.

Nigg, J., Carte, E., Hinshaw, S., & Treuting, J. (1998). Neuropsychological correlates of childhood Attention-Deficit/Hyperactivity Disorder: Explainable by comorbid disruptive behavior or reading problems? *Journal of Abnormal Psychology, 107*(3), 468-480.

Ostrander, R., Weinfurt, K. P., Yarnold, P. R., & August, G. J. (1998). Diagnosing attention deficit disorders with the Behavioral Assessment System for Children and the Child Behavior Checklist: Test and construct validity analyses using optimal discriminant classification trees. *Journal of Consulting and Clinical Psychology, 66*(4), 660-672.

Piesco, S., Baker, D. B., Silva, P. A., & Brooke, M. (1996). Behavioral distinctions in children with reading disabilities and/or ADHD. *Journal of the American Academy of Child & Adolescent Psychiatry, 35*(11), 1477-1484.

Purvis, K. L., & Tannock, R. (1997). Language abilities in children with attention deficit hyperactivity disorder, reading disorder, and normal controls. *Journal of Abnormal Child Psychology, 25*(2), 133-144.

Robins, P. (1992). A comparison of behavioral and attentional functioning in children diagnosed as hyperactive or learning-disabled. *Journal of Abnormal Child Psychology, 20*(1), 65-82.

Rucklidge, J. J., & Tannock, R. (2001). Psychiatric, psychosocial and cognitive functioning of female adolescents with ADHD. *Journal of the American Academy of Child & Adolescent Psychiatry, 40*(5), 530-540.

Sharp, W. S., Walter, J. M., Marsh, W. L., Ritchie, G. F., Hamburger, S. D., & Castellanos, F. X. (1999). ADHD in girls: clinical comparability of a research sample. *Journal of the American Academy of Child & Adolescent Psychiatry, 38*(1), 40-47.

Shaywitz, B. A., Fletcher, J. M., & Shaywitz, S. E. (1995). Defining and classifying learning disabilities and attention-deficit/hyperactivity disorder. *Journal of Child Neurology, 10*(Supplement 1), S50-S57.

Steingard, R., Biederman, J., Doyle, A., & Sprich-Buckminster, S. (1992). Psychiatric comorbidity in attention deficit disorder: Impact on the interpretation of Child Behavior Checklist results. *Journal of the American Academy of Child & Adolescent Psychiatry, 31*(3), 449-454.

Note: The Authors wish to thank Anne Kuo for her capable help in the collection and coding of participant chart information.

Narcissism and Aggression: Is Inflated Self-Esteem Related to Aggressive Driving?

George E. Schreer
Manhattanville College

The purpose of this research was to examine the relationship between narcissism and aggressive driving behaviors. Based on the theory of threatened egotism, it was hypothesized that individuals with inflated views of the self that are challenged by another person would report higher levels of belligerent driving behavior. For this study, 91 participants (63 female, 28 male) completed self-report measures assessing narcissism, self-esteem, and aggressive driving behavior. In support of the theory of threatened egotism, specific aspects of narcissism (i.e., Exhibitionism for women and Entitlement for men) did predict aggressive driving behavior. Overall, these results suggest that inflated self-esteem, and not low self-esteem, may be an important cause of aggressive driving behavior.

The traditional belief among psychologists is that low self-esteem is a major predictor of violence and aggression (Levin & McDevitt, 1993; Toch, 1993). According to the low self-esteem theory, people are provoked by their own self-loathing to act out violently toward others or seek out situations in which their own self-worth is challenged in order to enhance their low self-esteem (Toch, 1993). Unfortunately, laboratory evidence supporting this direct link between low self-esteem and aggression is lacking (see Baumeister, Smart, & Boden, 1996, for an extensive review that questions this theory).

As an alternative to the low self-esteem theory, researchers have proposed that aggression may stem from threatened egotism (e.g., Baumeister, Bushman, & Campbell, 2000; Bushman & Baumeister, 1998; Papps & O'Carroll, 1998). According to this theory, people holding unrealistically highly favorable views of the self (i.e., high in narcissism) that are disputed or undermined by another person (i.e., ego threat) will be more prone to display aggression in order to defend their grandiose views of themselves (Baumeister et al., 2000). In support of their theory, Bushman and Baumeister (1998) found that people scoring high on narcissism displayed higher levels of aggression (as measured by blasts of noise) toward a person who allegedly criticized an essay they wrote compared to people who scored low on narcissism. In addition,

---

Author info: Correspondence should be sent to: George Schreer, Dept. of Psychology, Manhattanville College, 2900 Purchase Street, Purchase, NY 10577; email: schreerg@mville.edu.

North American Journal of Psychology, 2002, Vol. 4, No. 3, 333-342.

© NAJP

level of self-esteem by itself had no effect on aggression. Thus, contrary to the low self-esteem theory, the empirical evidence "best fits the view that aggression is most likely when people with a narcissistically inflated view of their own personal superiority encounter someone who explicitly disputes that opinion" (Baumeister et al., 2000, p. 28).

Threatened egotism has been used to explain many forms of violence including murders, rapes, and gang activity (Baumeister et al., 1996). Another form of aggression that may relate to the threatened egotism model is aggressive driving. In this context, a narcissistic driver may perceive another driver's action as a personal threat to their inflated view of themselves or their personal safety, and this may lead to retaliation in the form of belligerent driving behavior. As Deffenbacher asserted, "their car is their world, so if you threaten their car, you threaten them" (1999, as cited in Cornelius, p. B6). According to a recent survey, aggressive driving incidents are on the rise (AAA, 1997). It has been estimated that drivers may be subjected to hundreds of anger episodes and aggressive behaviors per year (Deffenbacher, Huff, Lynch, Oetting, & Salvatore, 2000). Based on daily driving records over a 10-day period, Neighbors, Vietor, and Knee (2002) found that participants reported having driving anger an average of 2-5 minutes per day.

Researchers investigating the causes of aggressive driving have typically focused on either situational determinants, such as the presence of aggressive stimuli (e.g., gun rack, aggressive bumper sticker; see Turner, Layton, & Simons, 1975) and anonymity (Ellison, Govern, Petri, & Figler, 1995). Sometimes they have focused on personality traits such as Type A behavior pattern (Elander, West, & French, 1993) and trait driving anger (Deffenbacher et al., 2000). For example, based on the driving data of college students collected over three days, Deffenbacher et al. (2000) found that students (especially males) that scored high on trait driving anger argued and fought more with other drivers.

Recently, researchers have begun looking into whether differences in motivational orientation may be associated with aggressive driving behavior. For example, Knee, Neighbors, and Vietor (2001) found a strong connection between aggressive driving and an ego-involved "controlled" personality type. Apparently, individuals with a controlled orientation were more likely to perceive the actions of other drivers as personal affronts to their self-esteem, and this subsequently led to more anger and aggression while driving. This new line of research nicely compliments the research done on threatened egotism. Much like people with an externally controlled personality type, narcissistic people are also sensitive to personal threats and more likely to approach events with their "self-esteem on the line."

The purpose of this study was to examine whether narcissism (and not self-esteem level) is related to aggressive driving behavior. Aggressive driving was defined as any driving behavior that intentionally endangers others (Ellison-Potter, Bell, & Deffenbacher, 2001). Aggressive driving includes a range of behaviors from tailgating, obscene gestures, and flashing the high beams to the more extreme behaviors such as deliberately obstructing the path of other vehicles and intentionally trying to injure or assault another driver (i.e., road rage). Risk-taking behaviors such as speeding and traffic weaving were not included in the definition. Narcissism is a pattern of grandiosity used to bolster and enhance a fragile sense of self-esteem. The Narcissistic Personality Inventory (NPI; Raskin & Hall, 1979) is the most widely used measure of narcissism. Using factor analysis, Raskin and Terry (1988) found that the NPI contained seven factors: Authority, Self-Sufficiency, Superiority, Exhibitionism, Exploitativeness, Entitlement, and Vanity. In particular, the components of Exhibitionism, Exploitativeness, and Entitlement have consistently been associated with hostility, aggression, and defensive self-esteem (Emmons, 1987; Raskin, Novacek, & Hogan, 1991a; Raskin & Terry, 1988; Wink, 1991).

Based on previous research investigating narcissism, it was predicted that certain aspects of narcissism (and not self-esteem level) would be better predictors of aggressive driving behavior. Specifically, it was predicted that the Exhibition, Entitlement, and Exploitativeness factors would correlate highest with aggressive driving. Gender was included as a variable, because current research has found certain aspects of narcissism (i.e., exploitive tendencies and feelings of entitlement) may be less central to the construct of narcissism among females than among males (Tschanz, Morf, & Turner, 1998). Finally, a measure of driving anger was also included to examine the validity of the aggressive driving measure developed for this study.

METHOD

Participants

The participants included 99 undergraduates (68 female and 31 male) attending a small, ethnically diverse college outside a densely populated metropolitan area in the Northeast. The average age of the participants was 20.7 years (*SD* = 1.8). Four students reported they did not drive and another four participants did not fully complete the questionnaires, so their data were not used in the analyses. Thus, the final sample consisted of 91 students (63 females and 28 males). The majority (72%) reported that they drive every day, and they drove a median of 60 minutes a day.

Measures

Self-Esteem. The Rosenberg Self-Esteem Scale (RSES, Rosenberg, 1965) was used as a measure of global self-esteem. Participants indicated their agreement with 10 items (e.g., "On the whole, I am satisfied with myself") using a 4-point scale (0 = strongly disagree, 3 = strongly agree). In the present sample, the alpha coefficient was .88.

Narcissism. The Narcissistic Personality Inventory (NPI; Raskin & Terry, 1988) was used to measure narcissism in a nonclinical population. The scale contains 40 items that are answered using a forced-choice (true-false) format. The NPI provides a full-scale score and includes seven components identified as Authority (e.g., "I like to have authority over other people"), Self-sufficiency (e.g., "I am more capable than other people"), Superiority (e.g., "I am an extraordinary person"), Exhibitionism (e.g., "I really like to be the center of attention"), Exploitativeness (e.g., "I can make anybody believe anything that I want them to"), Entitlement (e.g., "I insist upon getting the respect that is due me"), and Vanity (e.g., "I like to look at my body"). In the present sample, the NPI had an alpha coefficient of .79.

Driving anger. Driving anger was measured with the short form of the Driving Anger Scale (DAS; Deffenbacher, Oetting, & Lynch, 1994). Participants were are asked to rate the amount of anger that they would feel to 14 different driving situations (e.g., "Someone backs right out in front of you without looking"). Ratings were made using a 5-point scale from 1 (no anger) to 5 (very much anger). In the present sample, the DAS had an alpha coefficient of .87.

Aggressive driving. A 12-item aggressive driving questionnaire was created to measure behaviors often identified as indicators of aggressive driving (Knee et al., 2001; Neighbors et al., 2002). Similar to Knee et al. (2001), participants were asked to indicate how often they engage in various aggressive behaviors (e.g., tailgating, cursing, making obscene gestures) when confronted with another driver who makes them angry. Participants indicated their agreement with each item using a 5-point scale from 1 (never) to 5 (almost always). A reliability analysis found the full scale had good internal consistency (α = 80). For exploratory purposes, three additional items were included: "Do you feel vengeful if someone does something on the road to make you mad?," "Are you aggressive with other drivers if you cannot 'get back at' the driver who made you angry", and "Do you stay angry after a driver angers you?"

Procedure

After completing a consent form, the participants were given a packet containing the aggressive driving questionnaire, DAS, NPI, and self-esteem measures. In addition, demographic variables including age, gender, and driving habits (e.g., time spent driving per day, days spent driving per week, type of vehicle, and how long they had their license) were assessed. After completing the packet, the participants were then debriefed and thanked for their participation.

RESULTS

In support of the validity of the aggressive driving measure, aggressive driving scores correlated with driving anger ($r = .31$, $p < .003$), feeling vengeful if another driver makes you mad ($r = .56$, $p < .001$), and becoming aggressive with other drivers if you cannot get back at the driver who angered you ($r = .33$, $p < .001$). There were no gender differences in general self-esteem (*M*s = 31.11 for females and 31.07 for males) or in overall belligerent driving behavior (*M*s = 24.87 for females and 23.89 for males).

TABLE 1 Regression Analyses on Aggressive Driving Behavior as a Function of Self-Esteem and NPI Subscales for Females and Males

	Females (N = 61)		Males (N = 28)	
Predictors	β	t	β	t
Self-Esteem	.05	.38	.08	.26
NPI (Authority)	-.14	-.89	-.01	-.04
NPI (Exhibition)	.61	4.55**	.15	.56
NPI (Superiority)	-.25	-1.91	-.23	-.89
NPI (Entitlement)	.03	.25	.53	2.11*
NPI (Exploitativeness)	.07	.55	-.35	-1.23
NPI (Self-sufficiency)	-.04	-.29	-.01	-.03
NPI (Vanity)	-.22	-1.83	.04	.15

Note: $R^2 = .45$ for females, $R^2 = .32$ for males. *$p < .05$. **$p < .001$.

Consistent with the narcissism literature, males scored higher on the NPI than females (*M*s = 17.54 and 14.61, respectively), $t_{(87)} = 2.26$, $p < .03$. There were no gender differences on any of the NPI subscales with the exception that males scored significantly higher on the Entitlement factor than females, $t_{(88)} = -3.64$, $p < .001$. The NPI full scale was not associated with aggressive driving behavior; however, additional analyses were performed in order to examine whether certain subscales of the NPI were associated with aggressive driving behavior. Multiple regression analysis was conducted with aggressive driving behavior as the criterion and self-esteem and the seven subscales of the NPI as the predictors. All variables were entered simultaneously. As predicted self-esteem alone did not predict aggressive driving behavior ($\beta = .10$, $p < .39$), however, the Exhibitionism and Superiority subscales did predict aggressive driving behavior. Specifically, individuals scoring higher on Exhibition reported more belligerent driving behavior ($\beta = .46$, $p < .001$), and individuals scoring higher on superiority reported less aggressive driving behaviors ($\beta = -.28$, $p < .02$). In addition, analysis using partial correlations also found these associations held even when self-esteem was partialled out.

To investigate for gender differences, regression analyses were performed separately for males and females (see Table 1). For females, aggressive driving behavior was strongly predicted by Exhibitionism ($\beta = .61$, $p < .001$), and marginally predicted by Superiority ($\beta = -.25$, $p < .07$) and Vanity ($\beta = -.22$, $p < .08$). For males, however, the only significant predictor of aggressive driving behavior was the Entitlement subscale ($\beta = .53$, $p < .05$). The Exploitativeness subscale did not predict aggressive driving behavior for either males or females.

DISCUSSION

Consistent with current research (Baumeister et al., 2000), inflated self-esteem (i.e. narcissism) was a better predictor of aggressive driving behavior than low self-esteem. According to the theory of threatened egotism, people holding inflated views of themselves are more likely to display aggression when those grandiose views are threatened by another. These findings dovetail with current research investigating the relationship between motivational orientation and aggressive driving (Knee et al., 2001; Neighbors et al., 2002). Similar to individuals with a controlled orientation, narcissistic people may be more likely to perceive common roadway encounters as personal affronts to their self-esteem. In order to defend their strongly held views about themselves, narcissistic drivers may be more likely to retaliate in the form of belligerent driving behavior (e.g., cursing, yelling, honking, gesturing, and tailgating).

The NPI full scale did not predict aggressive driving behavior. Instead, researchers have proposed a two-component model where certain aspects of narcissism may be more indicative of nondefensive (healthy) self-esteem (e.g., Authority, Self-Sufficiency, Vanity, and Superiority), while other aspects of narcissism (e.g., Exhibitionism, Exploitativeness, and Entitlement) may be associated with defensive self-esteem and aggression (Emmons, 1987; Raskin, Novacek, & Hogan, 1991a; 1991b). In support of this two-component model, individuals scoring higher on Exhibitionism reported higher levels of belligerent driving behavior, whereas individuals scoring higher on Superiority reported lower levels of aggressive driving behavior. Contrary to expectations, Exploitativeness did not predict aggressive driving behavior. Perhaps Exploitativeness, as defined by Raskin and Terry (1988), did not predict aggressive driving behavior because these items may have been misclassified. For example, in another factor analysis on the NPI performed by Emmons (1987), the majority of items identified as Exploitativeness by Raskin and Terry (1988) fell into the less defensive category of Superiority/Arrogance.

Because narcissism may not describe the same phenomenon in both genders (Morf, Weir, & Davidov, 2000), gender differences were investigated. For females, aggressive driving behavior was predicted by higher Exhibitionism scores, and to a lesser extent, lower scores on the more nondefensive aspects of narcissism: Superiority and Vanity. For males, however, aggressive driving behavior was only predicted by higher scores on Entitlement. Thus, a strong sense of entitlement appears to be more characteristic of the male aggressive driver, whereas an excessive need for admiration may be more characteristic of the female aggressive driver. In partial support of these findings, Tschanz et al., (1998) found that feelings of entitlement were less central to the construct of narcissism among females than for males because such behaviors may "violate culturally held expectations regarding appropriate female behavior" (p. 868). Perhaps, women who engage in aggressive driving have a need to be noticed and they ensure this by yelling and making obscene gestures at other drivers. For male aggressive drivers, however, the act of driving may evoke a strong sense of power and privilege, and when that power is challenged, they react with hostility and aggression. Additional research should be done in order to elucidate whether different aspects of narcissism are more dangerous for one gender or the other.

Interestingly, women in this study reported the same levels of aggressive driving behavior as men. Much of the experimental research, however, shows that men are more aggressive than women and that the male gender role typically includes norms encouraging aggression (e.g., Eagly & Steffen, 1986). To explain this inconsistency, one needs to take into consideration the moderator of provocation. Although unprovoked men have been found to be more aggressive than women, provocation (e.g., frustration, insults, blows to one's self-esteem) greatly attenuates this gender difference (Bettencourt & Miller, 1996). According to Bettencourt and Miller (1996), provocation provides justification for aggression and frees women from the constraints evoked by gender role norms. In a typical aggressive driving situation, the driver is often provoked and angered by another, thus it is not that surprising that women displayed levels of aggressive driving behavior comparable to that of men. Future research could investigate whether situational factors such as provocation intensity might moderate gender differences in aggressive driving behavior.

In this study, the results are limited primarily to 18-22 year old college students. Community samples with a wide range of ages are needed to see how these findings generalize to other populations. In addition, to build on this research it may be useful to supplement the self-report measures with other methodologies. For example, having participants keep a diary of driving behaviors has the benefit of assessing actual events as they are experienced rather than studying imagined scenarios assessing how participants would respond (Neighbors et al., 2002). As an alternative, participants could partake in a driving simulator task (Ellison-Potter et al., 2001), where they are randomly assigned to high and low personal threat conditions. One could hypothesize that highly narcissistic people (especially Exhibitionism and Entitlement) will display more aggressive driving behavior when their inflated self-views are threatened.

REFERENCES

American Automobile Association. (1997). *Aggressive driving: Three studies.* (Washington, DC: American Automobile Association Foundation for Traffic Safety).

Baumeister, R. F., Bushman, B. J., & Campbell, W. K. (2000). Self-esteem, narcissism, and aggression: Does violence result from low self-esteem or from threatened egotism? *Current Directions in Psychological Science, 9,* 26-29.

Baumeister, R., Smart, L., & Boden, J. (1996). Relation of threatened egotism to violence and aggression: The dark side of high self-esteem. *Psychological Review, 103,* 5-33.

Bettencourt, B. A., & Miller, N. (1996). Gender differences in aggression as a function of provocation: A meta-analysis. *Psychological Bulletin, 119,* 422-447.

Bushman, B., & Baumeister, R. (1998). Threatened egotism, narcissism, self-esteem, and direct and displaced aggression: Does self-love or self-hate lead to violence? *Journal of Personality and Social Psychology, 75,* 219-229.

Cornelius, C. (1999, July 9). CSU professor hopes to drive down road rage. *Denver Post,* p. B6.

Deffenbacher, J. L., Huff, M. E., Lynch, R. S., Oetting, E. R., & Salvatore, N. F. (2000). Characteristics and treatment of high-anger drivers. *Journal of Counseling Psychology, 47,* 5-17.

Deffenbacher, J. L., Oetting, E. R., & Lynch, R. S. (1994). Development of a driving aggression scale. *Psychological Reports, 74,* 83-91.

Eagly, A. H., & Steffen, V. J. (1986). Gender and aggressive behavior: A meta-analytic review of the social psychological literature. *Psychological Bulletin, 100,* 309-330.

Elander, J., West, R., & French, D. (1993). Behavioral correlates of individual difference in road-traffic crash risk: An examination of methods and findings. *Psychological Bulletin, 113*, 279-294.

Ellison, P. A., Govern, J. M., Petri, H. L., & Figler, M. H. (1995). Anonymity and aggressive driving behavior: A field study. *Journal of Social Behavior and Personality, 10*, 265-272.

Ellison-Potter, P., Bell, P., & Deffenbacher, J. (2001). The effects of trait driving anger, anonymity, and aggressive stimuli on aggressive driving behavior. *Journal of Applied Social Psychology, 31*, 431-443.

Emmons, R. A. (1987). Narcissism: Theory and measurement. *Journal of Personality and Social Psychology, 52*, 11-17.

Knee, C. R., Neighbors, C., & Vietor, N. A. (2001). Self-determination theory as a framework for understanding road rage. *Journal of Applied Social Psychology, 31*, 889-904.

Levin, J., & McDevitt, J. (1993). *Hate crimes: The rising tide of bigotry and bloodshed.* New York: Plenum Press.

Morf, C. C., Weir, C., & Davidov, M. (2000). Narcissism and intrinsic motivation: The role of goal congruence. *Journal of Experimental Social Psychology, 36*, 424-438.

Neighbors, C., Vietor, N. A., & Knee, C. R. (2002). A motivational model of driving anger and aggression. *Personality and Social Psychology Bulletin, 28*, 324-335.

Papps, B. P., & O'Carroll, R. E. (1998). Extremes of self-esteem and narcissism and the experience and expression of anger and aggression. *Aggressive Behavior, 24*, 421-438.

Raskin, R. N., & Hall, C. S. (1979). A narcissistic personality inventory. *Psychological Reports, 45*, 590.

Raskin, R., Novacek, J., & Hogan, R. (1991a). Narcissistic self-esteem management. *Journal of Personality and Social Psychology, 60*, 911-918.

Raskin, R., Novacek, J., & Hogan, R. (1991b). Narcissism, self-esteem, and defensive self-enhancement. *Journal of Personality, 59*, 19-38.

Raskin, R., & Terry, H. (1988). A principle-components analysis of the *Narcissistic Personality Inventory* and further evidence of its construct validation. *Journal of Personality and Social Psychology, 54*, 890-902.

Rosenberg, M. (1965). *Society and the adolescent self-image.* Princeton, NJ: Princeton University Press.

Toch, H. (1993). *Violent men: An inquiry into the psychology of violence.* Washington, DC: American Psychological Association (original work published in 1969).

Tschanz, B. T., Morf, C. C., & Turner, C. W. (1998). Gender differences in the structure of narcissism: A multi-sample analysis of the Narcissistic Personality Inventory. *Sex Roles, 38*, 863-870.

Turner, C. W., Layton, J. F., & Simons, L. S. (1975). Naturalistic studies of aggressive behavior: Aggressive stimuli, victim visibility, and horn honking. *Journal of Personality and Social Psychology, 31*, 1098-1107.

Wink, P. (1991). Two faces of narcissism. *Journal of Personality and Social Psychology, 61*, 590-597.

Author note: I am grateful to Karen McLeod and Catalina Suarez for their assistance in collecting and coding the data. I would also like to thank Craig Johnson, Michele Gelles, and three anonymous reviewers for their helpful comments on a draft of this article.

Positive Attitude Toward Dreams: Reliability and Stability of a Ten-item Scale

Michael Schredl, Constanze Brenner, & Christina Faul
Sleep laboratory, Central Institute of Mental Health, Mannheim, Germany

The present study of approximately fifty college students investigated the reliability and stability of a ten-item scale measuring positive attitude towards dreams. The results indicate that attitude towards dreams can be measured reliably and that it is quite stable over time, whereas dream recall frequency increased during the study period. The relationship of attitude towards dreams and recall frequency appears not as strong as previously reported when items with direct reference to dream recall are not included in the scale. Since the attitude variable show higher correlations to personality traits than dream recall frequency, it will be interesting to study the effect of this variable on the relationship between dream recall frequency and personality variables.

Several studies have demonstrated a strong relationship between positive attitude towards dreams and dream recall frequency (Bartnicki, 1997; Belicki, 1986; Cernovsky, 1984; Herman & Shows, 1984; Hill, Diemer & Heaton, 1997; Robbins & Tanck, 1988; Schredl, Nürnberg & Weiler, 1996;Violani et al., 1990), e.g., $r = .41$ ($p < .001$; $N = 336$; Hill, Diemer & Heaton, 1997). A closer examination of the scales used in this area of research, however, revealed that these scales often included items with direct reference to dream recall frequency, e.g., "I do not pay any attention to my dreams" (Cernovsky, 1984), "Have you ever speculated about the possible meaning of one of your dreams" (Robbins & Tanck, 1988), or "I think about my dreams during the day" (Bartnicki, 1997). It seems obvious that scales including such items are strongly related to the person's ability to recall dreams and her or his dream recall frequency. Schredl et al. (2002) have shown that correlation coefficients are considerably smaller if items with direct reference to dream recall are not included in the scale ($r = .158$, $p < .004$; $N = 444$). Since Schredl et al.'s scale consisted mostly of items which were negatively formulated, e.g., "It is not necessary to engage in dream work." the purpose of the present study was to construct a scale with items measuring positive attitude towards dreams. In addition to the reliability determined as interitem

---

Author info: Correspondence should be sent to: Michael Schredl, Sleep Laboratory, Central Institute of Mental Health, D-68072, Mannheim, Germany. e-mail: Schredl@as200.zi-mannheim.de.
North American Journal of Psychology, 2002, Vol. 4, No. 3, 343-346.
© NAJP

consistency, which was high for the scales used in previous studies (e.g., $r = .80$; Schredl, Nürnberg & Weiler, 1996), the stability of the attitude scale was investigated. It was hypothesized that attitude towards dreams – comparable to personality traits – is quite stable.

METHOD

For measuring dream recall frequency, a seven-point rating scale "How often do you recall your dreams recently (several months)? (0 = never to 6 = almost every morning)" was applied (Schredl, 2002). The retest reliability of the scale is high ($r = .83$; $N = 39$; mean interval of 70 days; Schredl, 2002). On the basis of previous scales, twenty positively worded items measuring attitude towards dreams and without direct reference to dream recall have been formulated, e.g., "I like dreaming," "A person who reflect about her/his dreams, learns a lot about herself/himself,"and "I like to know more about dreams." The scales used a five-point Likert format (1 = not at all to 5 = total agreement).

Overall, 56 psychology students (49 women, 7 men) participated in the study. The mean age was 22.4 ± 5.1 yrs. The participants received the questionnaire and were informed that an additional questionnaire would be handed out four weeks later. They were not told that the same questionnaire was administered twice. The second questionnaire was completed by 48 participants. Since the dream recall variable was measured on an ordinal level, non-parametric techniques, e.g., Spearman-Rank order correlations, were computed for the analyses.

RESULTS

The sum score of all 20 items was correlated with the single items for the purpose of item selection. The correlation coefficients of 10 items were above $r = .60$ ($p < .001$). The sum of these items was used in subsequent analyses. The interitem consistency (Cronbach's alpha) for the ten-item scale amounted to $r = .905$ ($p < .001$; $N = 56$). For the retest, the coefficient was comparable ($r = .892$, $p < .001$; $N = 48$).

TABLE 1 Questionnaire Data (first administration and retest; $N = 48$)

Variable	First test	Retest	Statistical test	$r\,(t_1 - t_2)$
Dream recall frequency	3.96 ± 1.54	4.35 ± 1.36	Sign-Rank = 111.5*	.63**
Attitude towards dreams	37.8 ± 7.2	37.5 ± 6.4	t-test: $t = -0.3$	.73**

*$p < .01$; **$p < .001$

The means and standard deviations of the variables of the first administration of the questionnaire and of the retest are depicted in Table One. Whereas the sum score of the attitude scale did not differ between the two measurement points, there was a significant increase in dream recall frequency. The stability of the four-week period was high for both measures (see Table 1). The correlation coefficients between the attitude scale and the dream recall frequency scale were r = .161 (p = .128; N = 56, first test) and r = .054 (p = .299; N = 48, retest).

DISCUSSION

Results of the present study indicate that a ten-item scale is sufficient to measure positive attitude towards dreams reliably. Extending previous research, it was shown that the sum score of the scale was quite stable over the four-week period. It will be interesting to study what factors might affect attitude towards dreams. An experimental manipulation in which the participants were informed about the benefits of attending to dreams (e.g., for enhancing creativity, for problem-solving strategies) yielded a small, yet insignificant effect (Rochlen et al., 1999). Significant dreams that help to solve personal problems or suggest an artistic idea (Kuiken & Sikora, 1993; Schredl, 2000) might alter the positive attitude toward dreams.

The enhancing effect of a single measurement on dream recall frequency, as found in the present study, has been reported in the literature (Halliday, 1992; Redfering & Keller, 1974). Despite the increase in dream recall frequency, the attitude towards dreams score was stable, an indication that the attitude variable is not easily manipulated (see above).

As shown in a previous study (Schredl et al., 2002), the correlation coefficient between dream recall frequency and attitude towards dreams is small if items with direct reference to one's own dream recall frequency are not included in the scale. In the present sample, the coefficient did not reach significance due to small sample size but was comparable to the value of the larger study (r = .158, p < .004; N = 444; Schredl et al., 2002). Taking into account the findings that attitude towards dreams is more closely related to personality traits such as openness to experience than dream recall frequency itself (r = .41 (p < .001) vs. r = .21 (p < .01); N = 336; Hill, Diemer & Heaton, 1997; r = .464 (p < .001) vs. r = .158, (p < .004); N = 444; Schredl et al., 2002), it may be concluded that dream recall frequency is more susceptible to state factors such as frequency of nocturnal awakenings (cf. Schredl et al., 1998), sleep duration (e.g., Taub, 1970) or volitional paying attention to dreams (e. g., Redfering & Keller, 1974; present data).

To summarize, attitude towards dreams can be measured reliably and is quite stable over time. The relationship to dream recall frequency is not as strong as previously thought if items with direct reference to dream recall are not included in the scale. Since the attitude variable showed high correlations to personality traits, it will be interesting to investigate whether this variable plays a role in the relationship between dream recall frequency and the factors which might affect dream recall frequency.

REFERENCES

Bartnicki, K. J. (1997). *An exploration of life experiences, personality traits and sleep habits in relation to dream recall and dream content.* Sacramento: Master Thesis, California State University.

Belicki, K. (1986). Recalling dreams: An examination of daily variation and individual differences. In J. Gackenbach (Ed.), *Sleep and dreams: a sourcebook* (pp. 187-206). New York: Garland.

Cernovsky, Z. Z. (1984). Dream recall and attitude toward dreams. *Perceptual and Motor Skills, 58*, 911-914.

Halliday, G. (1992). Effect of encouragement on dream recall. *Dreaming, 2*, 39-44.

Herman, S., & Shows, W. D. (1984). How often do adults recall their dreams? *International Journal of Aging and Human Development, 18*, 243-254.

Hill, C. E., Diemer, R. A., & Heaton, K. J. (1997). Dream interpretation sessions: who volunteers, who benefits, and what volunteer clients view as most and least helpful. *Journal of Counseling Psychology, 44*, 53-62.

Kuiken, D., & Sikora, S. (1993). The impact of dreams on waking thoughts and feelings. In A. Moffitt, M. Kramer, & R. Hoffmann (Eds.), *The functions of dreaming* (pp. 419-476). Albany: State University of New York Press.

Redfering, D. L., & Keller, J. N. (1974). Influence of differential instruction on the frequency of dream recall. *Journal of Clinical Psychology, 30*, 268-271.

Robbins, P. R., & Tanck, R. H. (1988). Interest in dreams and dream recall. *Perceptual and Motor Skills, 66*, 291-294.

Rochlen, A. B., Ligiero, D. P., Hill, C. E., & Heaton, K. J. (1999). Effect of training in dream recall and dream interpretation skills on dream recall, attitudes, and dream interpretation outcome. *Journal of Counseling Psychology, 46*, 27-34.

Schredl, M. (2000). The effect of dreams on waking life. *Sleep and Hypnosis,2*, 120-124.

Schredl, M. (2002). Messung der Traumerinnerung: siebenstufige Skala und Daten gesunder Personen. *Somnologie, 6*, 34-38.

Schredl, M., Wittmann, L., Ciric, P., & Götz, S. (2002). Factors of home dream recall: a structural equation model. *Journal of Sleep Research* (submitted).

Schredl, M., Nürnberg, C., & Weiler, S. (1996). Dream recall, attitude toward dreams, and personality. *Personality and Individual Differences, 20*, 613-618.

Schredl, M., Schäfer, G., Weber, B., & Heuser, I. (1998). Dreaming and insomnia: dream recall and dream content of patients with insomnia. *Journal of Sleep Research, 7*, 191-198.

Taub, J. M. (1970). Dream recall and content following extended sleep. *Perceptual and Motor Skills, 30*, 987-990.

Violani, C., Ippoliti, C., Doricchi, F., & De Gennaro, L. (1990). Correlates of individual differences in dream recall. *Sleep Research, 19*, 144.

The Impact of Sport Team Identification and Attributions of Ability and Effort on Spectators' Impressions of Athletic Performance

Daniel L. Wann, Brian L. Keenan, Steve Burnett, Jennifer Martin, Leslie Page, & Lori Smith
Murray State University

Christopher D. Lantz
Truman State University

Previous work has indicated that individuals give more positive evaluations of successful performance when the performance is attributed to effort than when the performance is attributed to ability. The current study tested this effect in a sport setting by hypothesizing a two-way interaction involving attribution type (effort versus ability) and degree of identification with the team and player in question (low versus moderate to high). It was expected that moderately to highly identified fans would give more positive ratings to a player being recruited by their team if the player's successes were described as a function of his effort. Lower ratings were expected if the player's successes were attributed to his ability. No differences in ratings as a function of attribution type were expected among fans low in team identification. To test the predictions, subjects watched a taped practice of a player who was described as being recruited by the target team. Some participants read that the player's successful performances were due to effort while others were informed that they were due to ability. After watching the player practice, the participants evaluated him. The results confirmed the hypotheses. Discussion includes the racial implications of the current findings.

The attributions of spectators have been a topic of interest to sport scientists for several decades. Hastorf and Cantril (1954) conducted one of the first empirical investigations of the attributions of spectators. In this study, examinations of Dartmouth University and Princeton University students' descriptions of a football game between the schools revealed a biased pattern of attributions. For example, while 25 percent of the Dartmouth supporters classified the game as "rough but fair," only 2 percent of the Princeton fans viewed the game in this way. In contrast, the Princeton fans saw the game as "rough and dirty." When asked if the Dartmouth players had intentionally injured one of Princeton's star

---

Author info: Correspondence should be sent to: Daniel L. Wann, Department of Psychology, Murray State University, Murray, KY 42071 or to danwann@msumusik.mursuky.edu via Internet.

North American Journal of Psychology, 2002, Vol. 4, No. 3, 347-354.

© NAJP

players, 10 percent of the Dartmouth supporters agreed with this statement while 55 percent of the Princeton fans reported that the injury was intentional.

Subsequent work confirmed Hastorf and Cantril's (1954) findings. Many of these studies indicate a self-serving bias (Miller & Ross, 1975), in which the spectators report external factors (e.g., biased referees, bad luck) to explain team failures while reporting internal factors (e.g., talented, high level of effort) to account for team successes (Lau, 1984; Mann, 1974). However, recent work indicates that the self-serving bias is not found among all spectators. Rather, only those with at least a moderate degree of team identification utilize this strategy (team identification is defined as the extent to which an individual feels a psychological connection to a team or athlete, see Wann, Melnick, Russell, & Pease, 2001). Because the psychological well-being of identified fans is related to their team's performances, these fans have the most to gain through biased beliefs about the causes of an outcome.

The biased attributions of highly identified fans were documented by Wann and Dolan (1994). In their study, college students met prior to a basketball game involving their university's men's team that their team either won or lost. The participants were asked to complete a measure assessing their level of identification with the home team and then watched the contest. At the conclusion of the game, the respondents completed a questionnaire assessing their attributions of the game's outcome. Both internal attributions and external attributions were assessed. As expected, although spectators formed internal attributions following a win and external attributions after a loss, these self-serving attributions were only exhibited by fans with at least a moderate degree of identification with the team. Subsequent research by Wann and Schrader (2000) found similarly biased patterns among identified fans for different attributional dimensions.

Another attributional pattern found in general psychological work that may have relevance for the causality estimates of sport spectators involves attributions for ability and effort. Some authors have suggested that when evaluators are presented with two targets with equally successful performances, raters will give higher ratings to persons whose success was described as the result of a high level of effort than to persons whose success was due to natural ability (Baron & Greenberg, 1990; Wann, 1997). For instance, Mitchell, Green, and Wood (1981) had participants act as supervisors for three confederate subordinates. Two of the confederates performed at an average level while the third confederate's performance was above standard. The participant supervisor was presented with information indicating that the target subordinate's good performance was due to either effort or ability. The participants were then given a chance to rate the subordinates. Ratings of the target subordinate revealed more positive evaluations when the performance as ascribed to effort than when it was attributed to ability.

Based on the aforementioned research, it was predicted that sport spectators would be more impressed by an athlete's positive performance if his or her performance was thought to be due to effort than if it was believed to be a function of ability. However, not all spectators were expected to exhibit this attributional bias. Rather, consistent with the work provided above (e.g., Wann & Dolan, 1994; Wann & Schrader, 2000), it was hypothesized that only those fans with at least a moderate degree of identification with a player's team (or potential team) would report differences based on attributions of effort and ability. For persons low in identification with a target team, the role of team follower is only a peripheral component of the self-concept, at best (Wann, Royalty, & Roberts, 2000). Consequently, as predicted by social identity theory (Tajfel, 1981), the team's (and players') positive performances would not have implications for the self-worth of these fans. Indeed, fans with a low degree of team identification rarely report strong affective responses to watching the team perform (Wann, Dolan, McGeorge, & Allison, 1994).

METHOD

Participants

Participants were 100 university students earning extra course credit in exchange for participation. Twenty-two participants were removed from the sample (see below), resulting in a final sample of 78 participants (24 male; 54 female). They had a mean age of 20.19 years (*SD* = 3.54, range = 18 to 42).

Materials and Procedure

Upon entering the testing room and providing their consent to participate, participants (tested in groups ranging in size from 6 to 31) were asked to watch a video of a Division II college basketball team conducting a practice. The participants were informed that the team was a high school team preparing for an up-coming tournament. Participants then watched a one-minute segment of the video to acquire a general understanding of the fast break drill shown on the tape. After previewing the video, participants were told that they would now be watching the video in its entirety (approximately 10 minutes). They were also told that, while watching the video, they were to focus their attention on one player in particular (the participants were told the player's uniform number and position, point guard). They were told that this player was a junior who was being recruited to play basketball for the University of Kentucky. The participants then randomly received and read one of two different descriptions of the target player. For the most part, the two descriptions were similar, indicating that the player was very skilled. For instance, both descriptions stated that the player had averaged 26 points per game and had led his team to the state championship. However, the descriptions differed with respect to the reasons presented for the player's success. One group of participants, labeled the Ability group, read that "the player had been good his whole life because of natural ability". The other group of subjects, the Effort group, read that the player "was good because he worked hard and displayed a high level of effort." These descriptions were designed to lead to differing attributions to account for the player's success. After reading the target player descriptions, the descriptions were collected and the participants watched the entire video. The target player attempted four shots during the fast break drill, making one. He also had two steals, two rebounds, and one assist. This level of performance was similar to the other players involved in the practice.[1]

At the completion of the video, the participants were asked to complete a questionnaire packet containing four sections. The first section contained 10 items assessing impressions of the target player. Sample items included "How much do you think the observed player will contribute to the University of Kentucky's success?" and "How much leadership do you think the observed player will bring to the University of Kentucky's basketball team?" Responses to the items were Likert-scale in format. Anchors ranged from 1 (e.g., *he will not contribute*) to 8 (e.g., *he will contribute a great deal*). Thus, higher numbers indicated a more positive impression. The second section contained manipulation check items asking the respondents to recall the team recruiting the player and why the player was good at basketball.

The third section of the packet contained demographic items assessing age and gender. The final section contained the *Sport Spectator Identification Scale* (SSIS), a reliable and valid instrument for assessing team identification (Wann & Branscombe, 1993). The SSIS contains 7 Likert-scale items with response options ranging from 1 (*low identification*) to 8 (*high identification*). A sample item from the SSIS reads, "How important to you is it that (name of target team) wins?" The respondents targeted the University of Kentucky men's basketball when completing the SSIS.

After the participants had completed their packet, they returned it to a researcher who handed them a debriefing statement. This statement disclosed the hypotheses of the study and contained information on contacting the researcher for a final project report. The sessions lasted approximately 20 minutes.

RESULTS

Manipulation Checks

Twenty participants were removed from the original sample because they incorrectly selected the attribution for the target player's success (i.e., they indicated that the player's success was due to effort when if fact it had been due to ability or vice versa). Two additional participants were removed from the sample because they incorrectly stated which school was actively recruiting the target player. Therefore, all subsequent analyses were conducted on the remaining 78 participants.

Perceptions of Performance

The ten items assessing perceptions of the target player's performance were combined to form a single index of perception (Cronbach's alpha = .78). The sum of the ten items was then divided by 10 to acquire a measure of perception that was consistent with the original parameters of the scale (i.e., scores could range from 1 to 8 with higher numbers indicating more positive impressions). The seven items comprising the SSIS were also combined to form a single index of identification (alpha = .97). A pair of one-way ANOVAs was used to test for gender differences in impressions of the target player and level of team identification. The ANOVA computed on perceptions of the target revealed no differences in impressions reported by male ($M = 4.50$, $SD = 1.43$) and female ($M = 5.04$, $SD = 1.36$) participants, $F_{(1, 77)} = 2.51$, $p > .05$. The ANOVA examining level of identification also failed to find differences for males ($M = 19.17$, $SD = 16.21$) and females ($M = 21.85$, $SD = 14.49$), $F(1, 77) = 0.57$, $p > .05$. Thus, all analyses were conducted across gender.

A median split was used to categorize participants into two groups: those with a low level of identification ($n = 37$, $M = 8.89$. $SD = 2.42$, range = 7 to 14) and those with a moderate to high level of identification ($n = 41$, $M = 31.98$. $SD = 11.81$, range = 15 to 53). The hypothesis that spectators will report more positive impressions of a target player when his success is described as the result of effort rather than ability but only if the participant is at least moderately identified with (i.e., cares about) the team in question was tested through a 2 (Level of Team Identification: moderate-high or low) x 2 (Attribution Type: effort or ability) between-subjects ANOVA.

The ANOVA failed to reveal a significant main effect for both team identification, $F_{(1, 74)} = 1.59$, $p > .21$, and attribution type, $F_{(1, 74)} = 0.65$, $p > .42$. However, the Team Identification by Attribution Type interaction was significant, $F_{(1, 74)} = 4.10$, $p < .05$. An examination of Table 1 indicates that the pattern of effects was consistent with expectations. Specifically, among those low in identification with the target team, there was little differentiation in target impressions among those who read that the athlete's success had been due to effort (n = 17) versus ability (n = 20). Conversely, for those persons with at least a moderate level of identification, ratings of the target player were much higher when the athlete's past success was reported to be due to effort (n = 24) than when the success was described as a result of ability (n = 17). A t-test performed on persons with a moderate-high level of team identification revealed a significant difference in attribution type, $t_{(39)} = 2.15, p < .05$. However, a similar analysis of those low in team identification was not statistically significant, $t_{(35)} = -0.83, p > .40$.

TABLE 1 Means and Standard Deviations for Impressions of the Target Player by Level of Team Identification and Attribution for Performance (effort versus ability)

	Attribution for Performance			
	Effort		*Ability*	
Level of Identification	*M*	*SD*	*M*	*SD*
Moderate-high identification	5.43	1.11	4.57	1.45
Low identification	4.43	1.22	4.83	1.66

DISCUSSION

The research described above provided clear support for the hypotheses as fans give more positive evaluations to an athlete whose success was attributed to effort than when his success was credited to his ability. This effect was limited to those fans with at least a moderate degree of identification with the team. Thus, in addition to other attributional errors previously identified in sport settings, such as the self-serving bias (Wann & Dolan, 1994; Wann & Schrader, 2000) and the fundamental attribution error (Wann, Brewer, & Carlson, 1998), we can now add the effort-ability bias. This type of bias had been previously noted in an organizational setting (Mitchell et al., 1981).

The current findings have a number of important implications for sport settings, particularly for coaches and others in positions involving the evaluation of talent and performance. The data suggest that when coaches are confronted by athletes with equally successful performances, the coaches will be more impressed by those athletes whose performances were attributed to high levels of effort. Consequently, athletes whose performance was thought to be due to their effort will likely receive more positive evaluations. The result could be that athletes with equally positive skills do not receive equally positive evaluations from their coaches, resulting in biased assessments. The data provided above suggest that this process would only occur for those coaches with at least a moderate degree of identification with their team and players.

It is also possible that an inverse relationship would occur for coaches evaluating players who have performed poorly. Mitchell et al. (1981) found that when a subordinate's poor performance was attributed to a lack of effort, the subordinate received lower evaluations than a subordinate whose similarly poor performance was attributed to ability. Thus, similar to athletes who succeed, athletes who perform poorly may not receive equal treatment if a moderately to highly identified coach assigns different causes (i.e., effort versus ability) to explain the athletes' failures (although the coach's status as head or assistant coach may impact this relationship, see Carver, DeGregorio, & Gillis, 1980). A lack of effort would be viewed more negatively than a lack of ability, a consequence of the controllability of the attribution (Wann & Schrader, 2000).

A final implication of the current study involves racial differences in attributions for athletic performance. Research indicates that the successful performances of African American athletes tend to be stereotypically attributed to the athletes' natural ability, while the successes of Caucasian athletes tend to be viewed as the result of high levels of effort (Harris, 1993; Murrell & Curtis, 1994). In light of the present study, these differential attributions could be quite problematic for minority athletes. That is, although African American and Caucasian athletes may be equally successful, because the white athletes' successes are believed to be a consequence of effort, these persons will receive more positive evaluations than black athletes. As Wann (1997) notes, the result of these biased attributions based on race presents "yet another example of the challenges facing minority athletes as they strive to compete on an even playing field" (p. 198).

REFERENCES

Baron, R. A., & Greenberg, J. (1990). *Behavior in organizations: Understanding and managing the human side of work* (3rd ed.). Needham Heights, MA: Allyn and Bacon.

Carver, C. S., DeGregorio, E., & Gillis, R. (1980). Field-study evidence of an ego-defensive bias in attribution among two categories of observers. *Personality and Social Psychology Bulletin, 6,* 44-50.

Harris, O. (1993). African-American predominance in collegiate sports. In D. Brooks & R. Althouse (Eds.), *Racism in college athletics: The African-American athlete's experience* (pp. 51-74). Morgantown, WV: Fitness Information Technology.

Hastorf, A. H., & Cantril, H. (1954). They saw a game: A case study. *Journal of Abnormal and Social Psychology, 49,* 129-134.

Lau, R. R. (1984). Dynamics of the attribution process. *Journal of Personality and Social Psychology, 46,* 1017-1028.

Mann, L. (1974). On being a sore loser: How fans react to their team's failure. *Australian Journal of Psychology, 26,* 37-47.

Miller, D. T., & Ross, M. (1975). Self-serving biases in the attribution of causality: Fact or fiction? *Psychological Bulletin, 82,* 213-225.

Mitchell, T. R., Green, S. G., & Wood, R. E. (1981). An attributional model of leadership and the poor performing subordinate: Development and validation. *Research in Organizational Behavior, 3,* 197-234.

Murrell, A. J., & Curtis, E. M. (1994). Causal attributions of performance for black and white quarterbacks in the NFL: A look at the sports pages. *Journal of Sport & Social Issues, 18,* 224-233.

Tajfel, H. (1981). *Human groups and social categories*. Cambridge, UK: Cambridge University Press.

Wann, D. L. (1997). *Sport psychology*. Upper Saddle River, NJ: Prentice Hall.

Wann, D. L., & Branscombe, N. R. (1993). Sports fans: Measuring degree of identification with the team. *International Journal of Sport Psychology, 24,* 1-17.

Wann, D. L., Brewer, K. R., & Carlson, J. D. (1998). Focus of attention and sport spectators: Beliefs about causation. *Perceptual and Motor Skills, 87,* 35-41

Wann, D. L., & Dolan, T. J. (1994). Attributions of highly identified sports spectators. *Journal of Social Psychology, 134,* 783-792.

Wann, D. L., Dolan, T. J., McGeorge, K. K., & Allison, J. A. (1994). Relationships between spectator identification and spectators' perceptions of influence, spectators' emotions, and competition outcome. *Journal of Sport and Exercise Psychology, 16,* 347-364.

Wann, D. L., Melnick, M. J., Russell, G. W., & Pease, D. G. (2001). *Sport fans: The psychology and social impact of spectators.* New York: Routledge Press.

Wann, D. L., Royalty, J., & Roberts, A. (2000). The self-presentation of sport fans: Investigating the importance of team identification and self-esteem. *Journal of Sport Behavior, 23,* 198-206.

Wann, D. L., & Schrader, M. P. (2000). Controllability and stability in the self-serving attributions of sport spectators. *Journal of Social Psychology, 140,* 160-168.

---

[1] It warrants mention that the experiment could have been completed without the video. That is, we could have simply asked subjects to read a description of a hypothetical player whose successful performance was due to either effort or ability. We chose to include the video to increase the believability of the study. Because the target player's performance did not stand out as being either particularly stellar or particularly poor, the video should merely have had the desired effects of increasing experimental realism.

AN INTERVIEW WITH ALBERT ELLIS ABOUT RATIONAL EMOTIVE BEHAVIOR THERAPY

Albert Ellis
Albert Ellis Institute, New York City

(interviewed on behalf of NAJP by)
Michael F. Shaughnessy
Eastern New Mexico University
and
Virginia Mahan
South Plains College

NAJP: Looking back over the last 50 years, what hath REBT wrought? What has it done, and what changes has it brought about in the field of therapy and counseling?

AE: In January 1955, when I first started to do REBT, after I figured it out from 1953 through 1955, there were practically no cognitive behavior therapies and the main cognitive therapy was Alfred Adler's therapy, but that had very few behavioral and emotive aspects. So REBT – Rational Emotive Behavior Therapy - was one of the pioneering cognitive behavior therapies, which was followed more than 10 years later, by Aaron Beck's Cognitive and Donald Meichenbaum's Cognitive Behavior Therapy. Since the 1970s, the cognitive behavior therapies have become very popular and many practitioners of psychoanalysis and other therapies often use its methods. So it has revolutionized the field of psychotherapy and it has produced many outcome studies, which shows that it works as well or better than other major forms of therapy. The field of therapy and counseling has changed enormously as a result of Rational Emotive Behavior Therapy and the Cognitive Behavior Therapies that followed it.

NAJP: Could you address the issue of culture and how it often works against success? I am thinking of our fast food, expediency, and immediate gratification culture.

Author info: Correspondence should be sent to: Dr. Michael Shaughnessy, Psychology Dept., Eastern New Mexico U., Portales, NM 88130.

North American Journal of Psychology, Vol. 4, No. 3. 355-366.
© NAJP

AE: According to REBT, it would be better if people practiced more long range hedonism and less short range hedonism and didn't mainly go for pleasures of the moment but also for the pleasures of the moment that also help them in the future. Our particular culture pushes fast foods, expediency, and immediate gratification. So it encourages short-range hedonism. This is because the economic system under which we work benefits more by short-range hedonism in some ways. Therefore, we could benefit by showing people how not to be hung up on immediate gratification and how to do better today and also in the future.

NAJP: What work still needs to be done in terms of REBT?

AE: REBT can help people to think more rationally, to feel less anxious, depressed, and enraged when they fail and get rejected, and to act more hopefully to get more of what they want and less of what they don't want.

NAJP: What are the main "shoulds," "oughts," and "musts" that we need to address in American culture, and in our modern day technological world?

AE: Rational Emotive Behavior Therapy uniquely says that much human disturbance is self-created by people saying 1) "I prefer to do well and be approved, and loved by others, and I absolutely have to do so!" 2) "Because I prefer you to treat me kindly and fairly, you absolutely have to act that way!" The first of these irrational beliefs leads to depression, despair and anxiety; and the second one to a great deal of anger and rage. The third common irrational belief is: "I prefer conditions in which I live to be nice, fair, and easy, and to give me what I want and not give me what I don't want; and, therefore, I need those conditions to be nice, fair, and favorable, or else it's all awful, terrible, and horrible, and I can't enjoy myself at all!" Now in our culture, we particularly favor these three absolutistic "musts," and because we have them, we may have more disturbed people than other cultures. Other cultures have dogmatic musts such as, "I must succeed not just in making money but in getting along with my family!" All cultures encourage the three main "musts," but maybe the American culture has them worse than some of the other cultures.

NAJP: What do you see as your biggest failure over the past 60 years?

AE: My biggest failure may be that I haven't convinced many therapists that in order to help people change for the better - that is, to get better and not only feel better - therapists preferably should better concentrate not on *only* cognitive, or *only* emotive, or *only* behavioral techniques of therapy. All three processes - cognition, emotion and behavior - interact to create human disturbances. Therefore, clients had better change their self-defeating thinking but do so emotionally and behaviorally. They had better change their feelings, but work at them cognitively and behaviorally. They had also better change their behavior, but do so cognitively and with strong feeling. Many therapies emphasize one of these three things, not thinking, and feeling, and action, as REBT does. REBT keeps using many cognitive, many emotional, and many behavioral methods. Yes, all three!

NAJP: What types of client have given you the most trouble, grief or aggravation? Which types of problems are most difficult to treat with REBT?

AE: As with other therapies, REBT clients with severe personality disorders and psychotic behaviors are more difficult to treat. We often treat them with REBT and sometimes we do remarkably well. But, they usually have serious biological elements (cognitive, emotional, and behavioral) in their make-up as well as socially learned characteristics. Therefore, they rigidly hold their absolutistic "shoulds" and "musts" and often won't work to give them up. They take longer to benefit as much from the therapy as our other neurotic clients. But since we have so many techniques, cognitive, emotive, and behavioral, if we keep working with severely disturbed clients, we can usually teach them, especially with homework assignments, to use several of our methods. But again, it takes longer and they don't get as improved as do the nice neurotic clients.

NAJP: You have frequently quoted Epictetus and other famous philosophers. At this stage of the game, what do you see as the single most important philosophical tenet that we need to adopt? Who would you say (other than yourself) is the most important philosopher that we need to read?

AE: First of all, REBT is more Epicurean than was Epictetus. Epicurus was all for engaging in pleasure. But he said that people better have it in a disciplined manner; and, therefore, he advocated long-range hedonism, rather than short-range hedonism. REBT follows many other philosophers, including modern ones like Immanuel Kant, Bertrand Russell, and John Dewey. It follows the existential philosophies of Heidegger and Sarte. It holds that people are constructivists and partly choose to have their own disturbed and non-disturbed thoughts, feelings, and behaviors. But, unlike some of the other constructivist therapies, which are passive, REBT active-directively teaches clients how to choose more functional solutions to their lives. Specifically, REBT shows clients how to work at achieving the philosophy of unconditional self-acceptance (USA), unconditional other-acceptance (UOA), and unconditional life-acceptance (ULA). To achieve this, they are encouraged to fully accept themselves and other people with their flaws and warts and to accept (but not like) unfortunate life conditions that they are unable to change. Like Adlerism, REBT strongly encourages clients to adopt more workable and less destructive philosophies of life.

NAJP: What irrational thoughts, if any, are unique to the later years of life? How would you best dispute them?

AE: Well, older people are more confronted with death and therefore frequently have the irrational thought, "Now that I am living, I must not die, and my loved ones must not die." So they're more anxious about dying. They also have many more handicaps than other people – illnesses, the lack of ability to get around, or the loss of their friends and relatives. So they are likely to have less tolerance of frustration. They have the usual irrational "musts" of younger people - "I must perform well!;" "Other people must treat me nicely!"; "Conditions must not be so hard!" And they may have them more frequently and stronger than younger people. So Emmet Velten and I have written a book, *Optimal Aging*, that shows older people how to use REBT. Also, older people are more prone to depression and inertia. But, using REBT, we push them, encourage them and teach them to anti-awfulize in spite of their old age and in spite of their restrictions. We help them to follow Reinhold Neibuhr's serenity prayer: "Give me the courage to change what I can change, the serenity to accept the things that I cannot change (including the hassles of older age) and the wisdom to know the difference."

NAJP: What do you see as the positive and negative sides of the changing role of women in a) the home b) the workplace and more specifically in the field of psychology over the last quarter of a century?

AE: Well, I think that women, partly because of Women's Liberation, have definitely helped themselves and have also been helped by some liberal men. There is now less prejudice against them. So they're doing much better in several professional fields, including psychology. There are more women in psychology and they get more status than before. That is very good. On the other hand, they now have taken on a lot of the male view that they have to perform well and have to be outstanding, the way men think they have to do. So they have a double-burden, especially if they're married and have children. The *must* to be good mothers and wives, and they *must* also at the same time be achieving! So they have the main *musts* that REBT points out, especially, "I have to do well in order to be a worth-while person!" Also, to some degree, the world may be rougher than it was for women today, so they may develop low frustration tolerance. So I think that women may benefit from their freer status if they don't take the achievement *musts* of our culture too seriously. They can try to do reasonably well at work and at motherhood, but then refuse to escalate their views to, "I have to do *marvelously* well in both areas or I'm no damn good!" They then make things even worse than they were before. So the fact that women have more opportunities is great – if they don't put *musts, shoulds and oughts* on themselves. If they learn REBT, they can function even with the double-burden of trying to do well at work and at marriage – if, again, they don't think that they absolutely have to do so.

NAJP: What is the most positive aspect of living to age 90?

AE: Well, I'm not quite 90, I'm 88 and will reasonably soon be 89. So, let's assume that I live to be 90, which I probably will. If so, I may *know* more and be more reasonable than I used to when I was young. Because I then foolishly thought *more* that I had to do well, that other people had to treat me nicely, and that world conditions must be easy for me. But now I'm seeing so much trouble and turmoil in the world, that I realize that *first* I can only do limited things, and that's too bad, because I would like to do more and accomplish more, but I don't have to do so. No matter how I behave, many people will dislike or ignore me – because, they jealously think I did too well or critically think I didn't do well enough. The world in many respects, now that I'm older, is more prone to irrationality, to war, to feuds, to terrorism, etc., and I don't like that. If anything, I think that people are more disturbed than they used to be and I certainly don't like that. But I think I've gained some degree of unconditional acceptance. Too bad things are the way they are . . . but they are. So I'm definitely living with my arthritis, my diabetes (which I've had for many years), and with wobbly legs that don't carry me as well as they used to do. So again, that is too bad, but I *accept* it. I wish like hell it weren't so, but I accept it as it is. I'd certainly prefer to be younger, healthier, and more agile like I used to be. I'd also prefer to be more sexually active. I'm definitely still sexually active, but not *as* sexually active as I used to be. Too bad – but not awful. I would prefer things to be better than they are for me as well as for other people, but I think that I've gained some degree of real acceptance over the years . . . and that's good!

NAJP: What have you personally done to combat the more negative aspects of aging?

AE: The main thing I've done is to accept the unfortunate things I can't change, and to accept that I am *lucky* to have lived as long as I have. I like living – and I hope I live to be 150, but let's face it, I won't. Yes, old age has its restrictions and troubles. You could say I'm lucky because I've done well and I've been acknowledged by many people. But that too, being acknowledged and being famous, has its hassles! Because now more people than ever want to speak with me, want to take my picture, and want autographs signed, and that has its hassles. But I largely accept what I cannot change. I still do try to change, I try to do better, and I keep up with my exercise, maintain my rigorous diet (which I've had for years), and conquer various other difficulties. I sorrow for my relatives and my friends who are no longer with me. But I don't take their deaths or my coming demise too seriously. It's too bad that I'll die too, but I accept the fact that I eventually will. I am able to look upon the good side that I'm still alive and cheerful and can still do most of the things that I did before. As for what I *cannot* do right now - *too damned bad*! Not awful, not horrible, not terrible, just a pain in the ass. And I use my own philosophy, the REBT philosophy that I teach to others. Yes, things are tougher than they used to be when I was younger. *Tough shit! Too bad! That's just the way it is!* So I go on and enjoy what I can.

NAJP: How do you think increased exposure to the media has affected the average person's psychological functioning in light of the September 11th terrorism and anthrax bioterrorism?

AE: Media definitely catastrophize more and awfulize more than the grim facts are in order to sell newspapers, magazines and radio and TV shows. So they make things seem worse than they are. And they're bad enough, this time really bad. But people are prone to be suggestible and prone to listen to the awfulizing of the media. They take some things *too* seriously. Of course, there is maybe more terrorism and irrationality, but if we totally damn terrorists as people, we will be doing exactly what the terrorists are doing, hating us and deciding that we don't deserve to live. They deny us, we deny them, then they deny us back. Hatred of people begets more hatred of people – so damnation goes on and on! The solution? To teach *all* individuals from kindergarten onward to unconditionally accept other people and to hate only what they do. This, of course, is the Christian philosophy of damning the *sin* but not the *sinner*. REBT teaches clients – and other people – this kind of unconditional other–acceptance (UOA) – that is, always to accept people who you think are acting "wrongly" and through doing so to *unangrily* try to induce them to behave less "wrongly."

NAJP: Apply REBT to Clinton or Condit's widely published sexual escapades.

AE: REBT, in spite of the fact that I have written many liberal books on sex over the years, does not endorse any special sex customs. It usually recommends following the sex rules of the community in which you live because you often get into trouble if you don't. So it says if Clinton had adulterous relations and if he lied about them to the public and to his wife, that is wrong because the American community says it's wrong. And if he wants to live in America and follow his own sex rules, he's not a worm, or a louse, or a loser, or a bum for doing so, but we may condemn and punish him for what we consider his wrongdoing. Wrong is wrong in America, whether it is sexual or nonsexual. So Clinton, by our American standards, may well have acted badly, but that doesn't mean he's a rotten person, or even a rotten President. REBT judges what we did sexually in regular American terms – but it does not hold that his sexual errors make him a totally rotten individual. REBT supports fairly conventional morality, but it does not damn people who flout conventional rules. If you choose to live in America, you normally had better follow its main sexual and nonsexual laws and customs. But I don't rigidly and totally hold that you *absolutely must* follow them or else you are a worthless person. Therefore, you may put Clinton's sexual *behavior* down but refuse to damn him totally. Again, REBT says that social rules are made by communities and families. Therefore, you had better follow your community's rules and your family's rules. But, if you don't follow the rules, you are not a rotten person. If you think you are no good for doing what Clinton did, then how can a no-gooder like you change and do better in the future? "Right" and "Wrong" are cultural and community values. REBT doesn't say anything is absolutely right or wrong. But if you live in a certain culture that defines things as right and wrong, you'd better, in order to get along in that culture, accept the fact that your culture sees things as "good" or "bad". Otherwise, you'll be penalized for doing the wrong thing. But *that never makes you a bad person!!* It makes you a person who in your community did a "bad" act. Now what are you going to do in the future to do fewer "bad" acts and get less in trouble?

NAJP: What do you think is the biggest challenge in the field of psychology in the next decade?

AE: Let me talk about psychotherapy rather than psychology in general. Psychotherapy has done many outcome studies and practically every single one of them unfortunately tests whether clients *feel* better rather than *get* better. In REBT, we say that to get better you not only have to reduce, say, the anxiety that you suffer from, but also keep it minimal, and finally get to the stage where you practically never feel severely anxious (or depressed, or raging) about anything, even if the worst things happen to you. Now psychotherapy practically never tests whether you *get* better with treatment, but largely tests whether you *feel* better. Most clients feel better because they have a kind and caring therapist who likes them and listens to them. Not because they've learned basically sound therapeutic philosophy that will help them deal with themselves and other people for the rest of their lives. Little has been done to promote this aspect of psychotherapy; and, therefore, help clients to feel better, to get better, and to stay better. Psychotherapy could also be more oriented toward teaching people to have a more fulfilled and happier everyday life.

NAJP: How has medication influenced the field of psychotherapy and psychiatry?

AE: Some people are biologically or endogenously depressed or otherwise disturbed. In some cases, medication, if properly used, helps them to not only feel better, but to think better and get better. It may, therefore, aid their psychotherapy. On the other hand, medication is often abused and people only take it, instead of taking it plus psychotherapy. Therefore, sometimes it does more harm than good. On the whole, psychotherapists can understand the good that medication does as well as the questionable aspects of it. I personally often recommend that my severely depressed and panicked clients experiment with medication to see if it effectively works for them *as individuals*. They can then refrain from seeing it as a quick and easy cure-all. To assess the usefulness of their taking medication, my clients are seen by a reputable psychopharmacologist who experimentally determines which, if any, work best for him or her.

NAJP: Could you say a few parting words about the other leaders in psychotherapy, e.g. Aaron Beck, Lazarus, Meichenbaum, Carl Rogers, Viktor Frankl, Skinner, Perls, and any others that you feel deserve mention?

AE: Aaron Beck has contributed greatly, especially by research studies on cognitive behavior therapy. But I think his cognitive therapy is not as forceful and emotional-evocative as is REBT. However, it is a good form of therapy, along with the cognitive behavior therapy of Donald Meichenbaum who started to do it in 1969, long after I first used REBT. Don has added to and made some valuable additions to cognitive behavior therapy. I first turned on Arnold Lazarus to REBT in 1964 when he was mainly behavioral. He later developed multimodal therapy and showed that therapists had better have several kinds of techniques at their disposal and use them *all* selectively with their clients. Carl Rogers saw existentially, that people are OK no matter what they do and *gave* them unconditional, positive regard. But unfortunately he didn't *teach* it to them. He was too passive and thought that clients would model his unconditional positive regard for themselves. In REBT, as I have said, we actively *give* unconditional acceptance to clients but also *teach* them to give it to themselves. Viktor Frankl was an existentialist who helped clients achieve solid *meaning* in their lives. He also pioneered with paradoxical psychotherapy. Fritz Perls added some expressive-emotive techniques to therapy, but put down realistic and logical thinking that would help people solve their problems. The pioneer in the field, of course, was Sigmund Freud, who had some very good things to say, but over-stressed repression and the unconscious and did not have good behavioral or even emotive-evocative techniques. Alfred Adler was excellent with cognition and did marvels to help people revise their disordered thinking. But he used few behavioral and some emotional techniques. He was a pioneer; however, and was much more rational than the other leading therapists of his day.

NAJP: What is your view of all this psychological testing, e.g. the TAT, the MMPI-II, the Rorschach, the 16PF, and all these things?

AE: Well, they have little validity, especially the Rorschach and the TAT, which I consider largely a waste of time because they are built on psychoanalytic hypotheses that are not scientifically falsifiable. So I would rarely use them. Even pencil and paper personality tests of irrational beliefs are limited, because people can consciously say they do not subscribe to them when they really do. So testing peoples' irrational and rational beliefs is sometimes informative. But we have much to discover yet about the scores from such testing before we can say they are accurate.

NAJP: Let's face it. We all have a lot to do and not enough time. How can REBT help us deal with stress?

AE: Let me refer again to Epictetus and the Constructivist position. As he said and as other constructivists say, it is not the unfortunate things that happen to us - the stressors of our lives - that upset us, but our *view* of them. REBT enables people to experience stress and not greatly distress themselves about it, by telling themselves rational coping statements such as "I really don't prefer this stress I am experiencing, but I can probably reduce it, so I'll do my best to arrange that. But if I can't reduce it right now or later, I can accept it and live a less happy, but still enjoyable life. Too bad, but the stress won't kill me – and I can even learn some valuable things by experiencing it. If people use REBT to deal with stress this way, they will rarely make themselves very anxious or depressed.

NAJP: Divorce is still a major American concern. Does REBT offer any assistance for couples? And what is your take on the "relationship craze" that seems to be sweeping America, e.g. *Men are from Mars, Women are from Venus*, etc.?

AE: *Men are from Mars* and *Women are from Venus* and that sort of thing represent from what I can see almost sheer Bullshit! They are distinctions made up by some therapists who have no good evidence behind their allegations. REBT has specialized in marriage and relationship therapy for many years. It notably helps people in couples and families who are angry at each other. They prefer to have other family members behave differently and are sorry and disappointed when they don't. But when they are angry and vindictive, they insist that their partners absolutely *must* think and act "properly." This helps to create anger in the other family members and soon we have a very disruptive relationship. REBT usually is quite effective in marital and relationship therapy because it shows almost all of its clients how to achieve unconditional other-acceptance even when others in the relationship seriously disagree with them and behave "badly." It helps them see that they create their anger and that they also can uncreate it and agree to disagree. If family members follow it, many marriages can be saved. REBT does quite well with relationship therapy.

NAJP: Who would you say has most influenced you and your way of thinking?

AE: Again, philosophers more than psychologists. As I noted before, both ancient and modern philosophers helped me to see that many people can be shown how to think, feel and behave much better. So I read these philosophers and experimented with some of their ideas for myself, used them with my clients, and thus constructed REBT.

NAJP: What are you currently working on and doing?

AE: I'm doing my autobiography, which will probably take several volumes. I will try to be honest about myself, about my clients, about other people, and about my friends and relatives. I shall show how I have personally benefited from trying to follow a rational-emotive-behavioral philosophy of life, and then was more able to help my clients and readers. I am also revising some of my earlier books and bringing them up to date. I am especially showing that when therapists use cognitive techniques, they can do so in a strong, emotional manner, and that when they use emotional and behavioral techniques they can amalgamate them with hard-headed thinking. In this manner I am integrating REBT with some of the main other therapies. It has always been cognitive, emotive, and behavioral, but I keep trying to make it even more so.

NAJP: Norman Vincent Peale lived to be 98. Maybe you can beat him and make it to 100. That will be some "Roast." What immediate plans do you have for the future and would you care to now save some others some trouble and write your own obituary?

AE: Well my immediate plans are to keep working with individual and group therapy clients. I also shall keep doing my famous Friday night workshops where I interview clients in public and then get the audience to participate. I shall keep giving other workshops and training workshops all around the country and the world. I shall of course keep promoting Rational Emotive Behavior Therapy and other modes of cognitive behavior therapy. So, I intend to keep doing that. And my obituary would say: "When I was young I was very anxious about succeeding, especially when approaching young females. But by reading philosophy and later by practicing Rational Emotive Behavior Therapy on myself, I gave up my dire need to do well, as well as my dire need to be approved by others. I also stopped demanding that conditions in my environment had to be good. I finally got to the point where I could formulate and use REBT with my clients and workshop participants. But before I used it with others, I had to keep using it on myself. Now that I am 88, I can accept the fact that I will die in a few years, when I would like to life forever. Too damn bad! So, I'll be dead as a duck one of these years. I don't fear death and I don't fear dying. I hope to live in reasonably good health as long as possible, but if I get more pains and ailments and I don't live as well as I'd like to live, that's too damn bad. I don't *demand* that life be good, but I wish to hell it were. I don't upset myself any longer about anything; and, therefore, I enjoy life as much as I can enjoy it. I'm still limited – as I said before in this interview – physically, because of my diabetes and other ailments. I'm not as agile as I used to be. But I live with my limitations and enjoy myself as much as I can in spite of them. I think that makes for a good life. If I continue to do so, preferably until at least 100, and if I help others to do the same then I think I would have lived a very useful life. But I don't *have to* be useful. I don't *have to* be well known. I don't *have to* sell a million books. I don't *have to* promulgate REBT all over the world so that it becomes the most accepted kind of psychotherapy. These are just my strong *preferences*. That's my goal, that's my purpose - and I hope to achieve it. But it's not *necessary*, it's only *very desirable* and *preferential*. So I think I'll stick to those desirable and preferential goals for the rest of my life and thereby enjoy myself as much as I may. It was good speaking with you, Michael. I really enjoyed this interview.

NAJP: What questions have we neglected to ask?

AE: I think you've asked some very good questions and I have done my best to answer them. I can't think of any important ones that may have been omitted, but if I do I'll let you know!

Counselor-Client Matching on Ethnicity, Gender, and Language: Implications for Counseling School-aged Children

Julie Hall
Los Angeles Unified School District, CA

Dana Kaplan Guterman
Saugus Union School District, CA

Howard B. Lee
California State University, Northridge

Steven G. Little
State University of New York at Albany

This paper addresses the issue of counselor-client match in one to one therapeutic interventions (i.e., counseling). The primary objective of this investigation was to examine children and adolescents' counseling outcomes to determine if clients of different backgrounds would benefit from matching the client and counselor on ethnic, gender, and language factors. Multivariate analyses were performed on the independent variables of ethnicity, gender, and client-counselor language matches to determine if general psychological functioning and other variables were different between matched and non-matched groups. Overall, ethnicity, language, and gender matches led to improvements in treatment outcome. Implications for counseling practice are discussed.

In various parts of the country, and in particularly California, 2000 census data have shown a tremendous growth in cultural diversity in the general population (Barrett, 2001a, 2001b). As a result, a greater emphasis will be placed on mental health professionals in the coming years to provide culturally aware and sensitive services to the individuals they serve. However, research studies on investigating the efficacy of matching clients with therapists on key variables such as ethnicity, language, and gender have been sparse. Of those that have been done in the area of multicultural counseling, college students and adults were primarily used as participants (Sue, 1992). In addition, many previous

---

Author info: Correspondence should be sent to: Dr. Howard Lee, Psychology Dept., Cal State University, Northridge, 18111 Nordhoff St., Northridge, CA 01330.

North American Journal of Psychology, 2002, Vol. 4, No. 3, 367-380.

© NAJP

studies focused on only one ethnic group and their preference for specific counselor characteristics (Gim, Atkinson & Whiteley, 1990). Studies that have been done to date with respect to gender and ethnicity issues in counseling have been equivocal, showing a need for further research in this area (Russell, Fujino, et al., 1996). Some studies support and others negate the efficacy of counselor-client match on treatment outcome (Sue, 1992). The inconsistent findings observed might be a result of studying different populations, using a variety of methodologies, and other confounds of experimental research. Nonetheless, there is a plethora of research on matching characteristics between individuals in a counseling dyad and their relation to treatment, but very few using ethnicity and language.

Sue, Fujino, Hu, Takeuchi, and Zane (1991) used logistic regression to show ethnic match was a significant predictor of positive treatment outcome for Mexican-Americans. Ethnic and language matches were shown to be particularly important for those clients who were less acculturated to the Anglo culture and/or those whose primary language was not English. However, ethnic match appeared to be a far stronger predictor of length of treatment than it was of treatment outcomes. This finding may have been influenced by the fact that treatment duration is more easily operationalized and more standardized across different counselor-client dyads than is treatment outcome. Therefore, measures of treatment outcome may not have accurately reflected actual therapeutic differences. Similarly, Russell, Fujino, Sue, Cheung, and Snowden (1996), using the same data set described in Sue et al. (1991), found that Black-Americans, Mexican-Americans, and Asian-Americans had higher psychological functioning when the client and therapist were matched on ethnicity. Gamst, Dana, Der Karaberian, and Kramer (2000) found *Global Assessment Function* scores to be higher for Asian-Americans and Mexican Americans that were ethnically matched than Black-Americans and White-Americans who were matched.

Ethnic match to individuals less assimilated in American society has been shown to be an important factor in counseling relationships. However, ethnic match for these individuals is not always feasible due to the fact that counselors of different ethnicity are underrepresented in psychology and related fields. This may be the reason why Asian-Americans and Mexican-Americans severely underutilized services in Los Angeles County (Sue, 1992). As Sue (1992) noted, these ethnic groups are more likely to be immigrants with different cultures, varying perspectives on mental health care, and primary languages other than English. Anglo-Americans and Black-Americans are less likely to be immigrants and experience language barriers, making health services more accessible. In support of this, it has been found that Black-Americans over utilized services in comparison with the general population (Snowden & Cheung, 1990; Sue, 1977; Sue et al., 1991). Although Black-Americans over utilized services, Sue et al. (1991) found they also showed the highest drop-out rate from treatment after only one session, and they averaged fewer sessions overall. Those Asian-Americans who sought services averaged the most sessions and had the lowest dropout rates after only one session. A possible explanation could be that Asian-Americans are more likely to wait and seek services only when concerns have become severe and, therefore, attend sessions for a longer duration. Overall, Mexican-, Anglo-, and Asian Americans exhibited greater improvement on the *Global Assessment Scale* (Americans exhibited greater improvement on the GAS), a treatment outcome measure, than Black-Americans. This could be attributed to the fact that Black-Americans attended fewer sessions, and, therefore, received less treatment than other ethnic groups.

Bernstein, Hofmann, and Wade (1987) assessed participants' gender role, age, and race and found them to be particularly useful in predicting preferences for counselor gender. In a population of 169 university students (102 female and 67 male), between 53% and 70% of participants had clear preference for counselor gender, depending on the type of problem. The large majority of participants expressed either no preference or preference for male counselors for all problems except sexual difficulties. On sexual issues, same-sex preferences emerged. For vocational and academic concerns, older students and those who were white preferred male counselors. Black-American students preferred female counselors. Participants with personal intimate concerns who were androgynous and who were younger expressed no preference. Masculine sex typed individuals preferred male counselors, and older students preferred female counselors. When parental concerns were the issue, Black-American students preferred male counselors. Overall, these studies highlight the importance of looking at gender matches between counselor and client.

The above studies emphasize the need for further research investigating the counseling process in relation to client satisfaction. Due to the mixed findings and the lack of research in the area regarding the school age population, this study attempts to address the counseling needs of children and adolescents. Presently, this population is assigned to counselors according to staff availability with little regard to individual student preferences. However, findings indicate that certain characteristics of the counseling dyad may influence treatment outcomes. The findings from the Sue et al. 1991) and Russell et al. (1996) studies were important concerning counselor-client matching, but neither study specifically differentiates between adults and children clients. Therefore, the primary objective of the current investigation is to determine if the counseling outcomes of children and adolescents of different backgrounds would benefit from matching clients and counselor in salient variables such as ethnicity, gender, and language.

This study uses the same data previously analyzed in the Sue et al. (1991) and Russell et al. (1996) studies. This study specifically examines school-aged children and adolescents on three salient and potentially important variables in the counseling process: ethnicity, language, and gender matches. These analyses are compared to those found for adults by Sue, et al (1991) and Russell, et al (1996). The findings of this study could provide invaluable information for those mental health facilities, school systems, school psychologists, and school counselors with regard to the assignment of counseling caseloads.

METHOD

Participants

Participants were children and adolescent clients (6 to 17 years) seen in outpatient services in Los Angeles County mental health facilities between January 1983 and August 1988. The ethnicity of the therapist and client was determined through self-report. Chinese, Japanese, Korean, and Filipinos were grouped as Asian-Americans because each taken individually was smaller than the selected sample of Black American, Mexican American, and White American clients. Non-Mexican Latinos were excluded from the analyses as well as some Asian groups that have a very small representation. The final sample consisted of 1219 Black, 903 Asian, 1,498 Mexican, and 996 White clients (*N*=4,616).

Data Set

The data were obtained from the Automated Information System (AIS) used by the Los Angeles County Department of Mental Health Services (LACMHS). The primary uses for the AIS data were revenue collection, system management, clinical management, and research. Therapists completed a standardized form for every client seeking treatment and aid from county-funded mental health facilities. Data consisted of client demographic information, diagnostic and treatment variables, type and extent of treatment, and therapist background information. Although the reliability of the AIS data has not been evaluated, measures have been taken to ensure its accuracy. Further details on the data set can be found in Sue et al. (1991) and Russell, et al (1996).

Variables

The primary objective of the study was to examine the relationship between therapist-client ethnicity, gender, and client-therapist language match and global mental health functioning. The language match was only relevant to Asians and Mexicans.

Ethnic Match. When the therapist who rated the admission Global Assessment Scale (GAS) score was of the same ethnic background as that of the client a match was recorded as present (e.g., Asian-American' therapist - Asian-American client). If the ethnicity of the therapist and client was not the same, then a match was absent.

Gender Match. A gender match existed when the therapist who rated the admission Global Assessment Scale (GAS) score was of the same gender as that of the client (e.g., Male - Male and Female - Female). A gender match was either present or absent.

Language Match. A language match existed when the therapist who rated the admission Global Assessment Scale (GAS) score could speak the same language (dialect) as the client (e.g., Spanish - Spanish, Mandarin - Mandarin, Japanese - Japanese, etc.). A language match was either present or absent. The language match was only pertinent to Asians and Mexicans. A large majority of Blacks and Whites listed their primary language as English.

Variables Analyzed

As with the Sue, et al. (1991) and Russell, et al. (1996) studies, the *Global Assessment Scale* (GAS) score obtained at the time of the initial interview was one of the variables. The GAS score at the time of discharge was also used but its utility in this analysis is questionable. The counselor-client match or non-match is only available at the time of the initial interview. The GAS used by the LACMHS is similar to the *Global Assessment of Functioning* (GAF) Axis V rating of the *DSM-IV* (American Psychiatric Association, 1994). Therapists assign a GAS rating based on psychological, social, and occupational functioning. Scores range from 1 *(most severe impairment)* to 100 *(good general functioning).* Some researchers have found the GAS to have high reliability (Endicott, Spitzer, Fleiss, & Cohen, 1976) and good concurrent and predictive validity (Soelberg, 1989). Other variables analyzed included the total number days of treatment, the number of treatment sessions, and premature termination. Premature termination was defined as the failure to return for treatment after one session. We used the same definition as the one used by Sue, et al (1991), where getting the client to return for further treatment is very important in obtaining therapeutic benefits. Initial diagnosis of the child was also obtained at the time of the first interview. A summary of the diagnoses is given in Table 1.

Table 1: Admission Diagnosis

Diagnosis	%	Diagnosis	%
Mental Retardation	1.1	Attention Deficit Disorder	3.4
Development Disorder	0.3	Pervasive Dev. Disorder	0.4
Conduct Disorder	19.5	Anxiety	7.9
Other Disorder	10.9	Substance Abuse	0.1
Alcohol Abuse	0.2	Other Substance Abuse	1.4
Schizophrenia	1.1	Schizo-Form	0.5
Schizo-Affect	0.1	Other Psych.	0.9
Major Depression	3.5	Bipolar	0.3
Affect	10.1	Somatoform	0.3
Dissociative	0.1	Psychosexual	0.2
Impulse Control	0.5	Adjustment	36.2
Psychophysical	0.2	Personality	0.5
Unspecified	0.3		
Total		100.	

Data Analyses

Descriptive and inferential analyses were performed within each ethnic group and between ethnic groups. Regression analyses were used to examine the relationship between matching variables and criterion variables. Logistic regression was used for examining the effects of matching on dropout (premature termination) and multiple regression was used when the dependent variables were (1) the number of sessions and (2) the number of days in treatment. Where possible, the analyses of the children data were compared to the overall results obtained by Sue, et al (1991) on an adult sample. The current study also included analyses for GAS scores by gender. The Sue et al. (1991) study did not do this.

RESULTS

Since this study pertains to school-aged children and their associated mental health problems, Table 2 gives the descriptive statistics on children within each ethnic group. The percentage of gender matches and language matches across all ethnic groups obtained in the current study were similar to those found in Sue, et al (1991). The percentage of Mexican-Americans in the current study was higher (32.5 versus 21.9) than the percentage found in the Sue, et al (1991) study. Across all ethnic groups, the Sue, et al study had a higher percentage of females than the current study. Also, on the percentage of ethnic matches, Black Americans had a smaller percentage in the current study than the one in the Sue, et al study (22.2% versus 33.8%).

The Mexican-American children were the highest in using county mental health facilities, followed by Black-Americans, White-Americans, and Asian-Americans. These results differ from the Sue et al. study that had White-Americans using mental health facilities more than the other ethnic groups. In fact in the Sue et al. study, Mexican-Americans were the lowest group in utilizing mental health facilities in Los Angeles County. The current results indicated that Asian American children under-utilized the county facilities when compared to the other three ethnic groups. They also had the highest percentage of ethnic match and the lowest language match. However, their GAS scores were higher than the other groups. Asian-American children in the sample were also older than the children from the other ethnic groups. Black-Americans had the lowest

TABLE 2 Summary Statistics by Ethnic Group

Characteristics	White	Black	M-A	A-A	Pairwise Comparisons
Sample Size (%)	21.6	26.4	32.5	19.6	
% Female	43.6	34	40.3	42	
Ethnic Match %	72.3	22.2	26.9	29.8	
Gender Match %	56.1	53.9	54.8	54.6	n.s.
Language Match	91.0	92.8	86.5	74.8	
%	12.7	13.0	12.4	13.3	w<a, b>m, m<a
Age (Mean)	34.4	34.9	33.6	36.8	n.s.
GAS (A) Female	33.8	37.6	30.0	38.6	w<b, w>m, w<a
GAS (A) Male					b>m, m<a
	38.9	8.3	37.3	39.5	n.s.
GAS (D) Female	36.8	40.1	33.5	41.5	w<a, b>m, m<a
GAS (D) Male	34.1	36.7	31.4	37.8	w>m, w<a, b>m
GAS (A) Overall					m<a
	38.8	39.5	35.0	40.7	w>m, b>m, m<a
GAS (D) Overall	100.7	81.7	118.3	108.1	w>b, w<m, b<m
# Days in Treatment					b<a

M-A = Mexican-American

A-A = Asian-American

w = White, b = Black, m = Mexican, a = Asian; A = GAS at Admission, D = GAS at Discharge number of female children using the county facilities and the lowest percentage of ethnic match between client and therapist. The youngest group of children came from Mexican-American families. They also had the lowest GAS scores. A pattern apparent from Table 2 is the gender differences. Male children in the Black-American and Asian American groups received higher GAS scores than female children. The opposite pattern existed for Mexican-Americans and Whites.

Separate one-way analysis-of-variances between ethnicity were done following a significant one-way multivariate analysis of variance. These analyses were statistically significant at the .05 level. Table 2 shows the results of the Tukey HSD pairwise mean difference comparisons between the ethnic groups. On mean age differences, Asians were the oldest followed by Blacks, Whites and Mexicans. Mexicans were significantly younger than Asians and Blacks. Both admission and discharge GAS scores for females were not significantly different. However for males on GAS admission scores, Whites were significantly lower than Blacks and Asians but significantly higher than Mexicans. In addition, Blacks and Asians were significantly higher than Mexicans. On discharge GAS scores, White males ($M = 36.8$) were lower than Asian and Black males, but higher than Mexican males. Mexican males were significantly lower than Asian males. In both admission and discharge GAS scores, Asian males had the highest scores, followed by Blacks, Whites, and Mexicans.

In comparing the mean number of days in treatment, Mexicans had the highest number of days. Asians were second followed by Whites and Blacks. The difference between Blacks with Mexicans, Whites and Asians were statistically significant. Whites were significantly lower than Mexicans. All other comparisons were not statistically significant. The mean number of treatment days by ethnic group is given in Table 2.

For the total sample and for each ethnic group, the GAS Admit score was correlated with the GAS Discharge score (.81 to .85). The total number of days in treatment was correlated with the total number treatment sessions (.63 to .76). The correlation between GAS scores and treatment days and sessions was near zero. The admission and discharge GAS scores for children were consistently lower than the scores reported by Sue et al (1991).

Dropouts From Treatment

Dropout was defined as the failure to return for treatment after one session. There were ethnic differences contrary to the percentages Sue et al. (1991) found using adults. The percentages of dropouts were 14.4% for Black-Americans, 15.6% for Whites, 13.5% for Mexican-Americans, and 15.8% for Asian-Americans. In the Sue et al. study the percentages were 19.4%, 15.3%, 14.6%, and 10.7% respectively. Using the test of significance between two proportions, no significant differences were found between any of the ethnic groups. Ethnic minority children-clients did not show a greater propensity to prematurely terminate treatment.

In order to examine the relationship between client-therapist match (ethnic and gender similarity) and termination, a logistic regression analysis was performed for each ethnic group. The criterion was dropout and predictor variables included: client gender, age, initial diagnosis, GAS score at admission, gender match, language match, and ethnic match. Table 3 shows the estimated effects and significance of each individual predictor variable, controlling for the effects of all other variables. The estimated effects for categorical predictors can be interpreted as the odds ratios of dropping out. Results indicated that for all ethnic groups except White-Americans, ethnic match resulted in substantially lower odds of dropping out than for the unmatched clients.

TABLE 3 Estimated Effects for Variable Predicting Dropout after One Session by Ethnicity

Characteristics	White	Black	M-A	A-A
Age	1.193 *	1.144 *	1.221 *	1.265*
Sex	1.179	.938	.961	1.009
Gender Match	.799	1.263	.966	1.443 **
Language Match	.809	.840	.774 *	.905
Ethnic Match	.732	.457 *	.643 *	.209 *
GAS[a] Admissions	.984 *	1.008	.995	.975 *
Diagnosis	.993	.997	1.002	.998

M-A = Mexican-American

A-A = Asian-American

[a] GAS = Global Assessment Scale

* p < .01, ** p < .05

Gender match was not a significant variable in determining dropout for any of the ethnic groups. Other significant predictors of lower probability of premature termination were age (being older) for all groups and having a higher initial GAS score for Whites and Asians. The overall results suggest that ethnic match is a very important consideration for Asian American, Mexican-American, and Black-American children-clients in predicting premature termination.

Mean Number of Sessions

There was evidence using the Tukey HSD test at $\alpha = .05$ that Whites differed significantly from Blacks and Mexicans on the mean number of treatment days. Blacks differed significantly from Mexicans and Asians. Mexican-Americans had the highest average, followed by Asians and whites. Black Americans had the lowest mean number of treatment days.

To examine the relationship between predictor variables and number of sessions, a multiple regression was performed for each group. Results are displayed in Table 4. For each ethnic group, older age was related to greater number of sessions. Ethnic match was a significant predictor for Mexicans and Asians only. Gender and language matches were associated with number of sessions only for Mexican Americans. Other significant predictors of more treatment sessions were being a girl (for Whites, Blacks, and Mexicans) and having a diagnosis of psychosis (for Blacks and Mexican children-clients). Initial GAS scores was not significant for any of the groups.

TABLE 4 Estimated Effects for Variable Predicting Number of Sessions by Ethnicity.

Characteristics	White	Black	M-A	A-A
Age	–.219 ***	–.235 ***	–.198 ***	–.139 ***
Sex	–.06 *	–.063 *	–.098 ***	.021
Gender Match	.054[b]	–.024	.053 *	.004
Language Match	.010	–.018	.065 **	–.001
Ethnic Match	.054[b]	–.063	.057 *	.095 **
GAS[a] Admissions	–.05	–.051	–.014	–.038
Diagnosis	–.057[b]	–.064 *	–.056*	–.020

M-A = Mexican-American

A-A = Asian-American

[a] GAS = Global Assessment Scale [b] These were significant at $p < .10$

* $p < .01$, ** $p < .05$

DISCUSSION

The results of this study indicate that ethnic, language, and gender matches in a counseling dyad are related to treatment effectiveness and important client variables. For ethnic, language, and gender matched pairs; treatment outcome was found to be rated significantly higher than for non-matched dyads. Overall, matching helped lead to improvements in psychological functioning scores and minimize drop out from treatment sessions. Ethnic, gender, and language matches also showed longer duration of treatment when compared to non-matched groups.

These results support the hypothesis that ethnic, language, and gender matches improve treatment effectiveness. Sue et al. (1991) and Russell et al. (1996) have documented this for adults. This study has found that it is also relevant to school aged children. However, there were some major differences found between the Sue et al. (1991) study and the current investigation. Sue et al. found significant differences in drop out across ethnic groups. The current study did not find such differences. Likewise when examining dropout and number of sessions in treatment, the current study found age differences while Sue et al. did not. The current investigation found older children more likely to dropout and have fewer treatment session. Sue et al. reported Asian-American adults to have the lowest dropout rate. In this study, Asian-American children had the highest dropout rate among the ethnic groups. In the child sample, Black-Americans showed lower dropout when ethnically matched with a therapist. This was not true in the adult sample. Concerning the number of treatment sessions, Mexican-American children stayed in therapy longer when they were matched to a therapist on ethnicity, gender, and language. Asian-American children seem to stay in treatment longer if they are matched ethnically only. Black-American children showed little differences in the number of treatment sessions when they were matched either on ethnicity or gender. Gender match and ethnic matched had a weak effect on the number of treatment sessions for White-Americans.

These findings seem particularly notable in that Goldberg and Tidwell's (1990) study indicated any race, ethnic group, or gender composition of counseling dyad was not directly related to client satisfaction with counseling. In our study, treatment outcomes, as measured by dropout and number of sessions, improved for many ethnic groups. If length of treatment can be taken as a measure of satisfaction, the results found here suggests higher client satisfaction with the counseling process when the clients were matched on either ethnicity, gender, and/or language. Additionally, Hess and Street (1991) reported that Mexican-American high school students did not rate Mexican-American and Anglo-American counselors differently. This points to the possibility of a discrepancy in client versus counselor ratings for treatment effectiveness. Therefore, further research should incorporate multiple measures of effectiveness ratings to develop a more comprehensive view of counseling dynamics.

Sue et al. (1991) report that ethnic match was related to longer treatment. The present findings show that children and adolescents who were ethnically matched also remained in treatment longer than those who were not matched. The above results imply that children and adolescents are likely to stay in treatment longer when they and their counselor share similar cultural characteristics. On the other hand, Sue's population did not demonstrate significant change in the number of treatment sessions for gender match, whereas, gender match in the current study was related to significant improvement in outcomes as demonstrated by dropout and number of treatment sessions. This finding suggests that children and adolescents benefit from having a therapist of the same gender. This appears to be consistent with children's developmental tendency to relate to same sex adults and peers.

One of the major problems encountered in using this data base for adults or children is what happens to the client after the initial consultation. The data only reports if there was an ethnic, gender, or language match at the initial consultation. There are no data to tell us whether the client received the same counselor/therapist on the subsequent sessions or even a counselor who was of the same ethnicity or gender. Hence, only the initial GAS scores would have direct interpretation. The discharge GAS score is suspect even though the data for the sample used in this study consistently showed children who were matched on ethnicity, language, and gender showed improvement in the GAS scores from the time of admission till discharge.

Further analysis using GAS scores did yield some interesting results. Surprisingly, male children tended to be rated lower on the GAS scale than female children when they were matched ethnically. In fact, males showed lower GAS scores in the matched condition than in the non-matched condition. The nonexperimental nature of this study makes it difficult to determine the reasons for this. We may be able to speculate that the therapist may have had greater unfulfilled expectations from male members of one's own ethnic group. At a more micro level, the White and Mexican-American males that are matched ethnically to a therapist showed the largest amount of decrease. However, regardless of the initial interview taken by the therapist, all children of each ethnicity and gender benefited from matching when the number of treatment days and sessions were considered.

Along these same lines, the analysis of Black- and Mexican-Americans showed a large decrease in admission GAS scores for the ethnically matched group than the unmatched group (35% for Black-Americans and 22% for Mexican-Americans). Males were evaluated much lower than females in the matched group. It would appear that Black-American and Mexican-American therapists are more critical toward male members of their own ethnic group. However, when we consider less subjective measures such as the number of treatment days and sessions, Black- and Mexican- American males showed significant improvement if they were in the matched group.

Another unexpected finding was that younger children were matched more often with counselors of the same ethnicity and language than older children. This could be attributed to the fact that language in younger bilingual children is not as well developed, resulting in the need for a counselor who can meet their language or cultural needs. Another possibility is that the range of ages in our study was small, and therefore, small differences were more likely to found significant.

Future research in the area might expand on this research by combining the matching characteristics of ethnicity, language, and gender versus simply looking at each match separately for discriminating variables. It would be interesting to examine the best possible combination of the various matched conditions and their implications for treatment outcome.

It seems logical that children receiving counseling services in public clinics and schools would also benefit from such an arrangement. Schools may want to provide culturally similar school psychologists and school counselors whenever possible in order to best meet their diverse students' needs. If this can not be accomplished, it seems reasonable to suggest that counselors and therapists who deal with school-aged children should try to minimize cultural differences between themselves and the children when providing counseling services. One way of doing this is to become as well educated about the culture as possible.

REFERENCES

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Barrett, B. (2001a, August 8). Census has residents in tune with rest of U.S. *Daily News*, Retrieved August 15, 2001, from http://www.daileynews.com/news/articles/0801/06/new01.asp

Barrett, B. (2001b, August 11). Valley's Mexican population booming. *Daily News*, pp. A1, A14.

Bernstein, B.L., Hofmann, B., & Wade, P. (1987). Preferences for counselor gender: Students' sex role, other characteristics, and type of problem. *Journal of Counseling Psychology, 34*, 20-26.

Endicott, J., Spitzer, R. L., Fleiss, J. L., & Cohen, J. (1976). The global assessment scale. *Archives of General Psychiatry, 33*, 766-771.

Gamst, G., Dana, R. H., Der Karaberian, A. & Kramer, T. (2000). Ethnic match and client ethnicity effects on global assessment and visitation. *Journal of Counseling Psychology, 28*, 547-564.

Gim, R.H., Atkinson, D. R., & Whiteley, S. (1990). Asian-American acculturation, severity of concerns, and willingness to see a counselor. *Journal of Counseling Psychology, 37*, 281-285.

Hess, R. S., & Street, E. M. (1991). The effect of acculturation on the relationship of counselor ethnicity and client ratings. *Journal of Counseling Psychology, 38*, 71-75.

Russell, G. L., Fujino, D. C., Sue, S., Cheung, M-K., & Snowden, L. R. (1996). The effects of therapist-client ethnic match in the assessment of mental health functioning. *Journal of Cross-Cultural Psychology, 27*, 598-615.

Soelberg, S. (1989). There's more in the number than you think: New validity data for the Global Assessment Scale. *Psychological Reports, 64*, 455-461.

Sue, S. (1992). Ethnicity and mental health: Research and policy issues. *Journal of Social Issues, 48*, 187-205.

Sue, S., Fujino, D., Hu, L., Takeuchi, D., & Zane, N. (1991). Community mental health services for ethnic minority groups: A test of the cultural responsiveness hypothesis. *Journal of Clinical and Consulting Psychology, 59*, 533-540.

Parent Discipline Scale: Discipline Choice as a function of Transgression Type

N. L. Lopez
University of Michigan, Ann Arbor

H. G. Schneider
Appalachian State University

C. S. Dula
Virginia Polytechnic Institute and State University

The *Parent Discipline Scale* (PDS) was designed to measure the effects of children's transgressions on parental disciplinary choices. The items for the final version of the PDS represent four types of child behavioral transgressions: aggressive and non-aggressive social violations, and aggressive and non-aggressive safety violations. Parents (n =101) were presented with these items and were asked to endorse the likelihood of using three different types of discipline techniques: non-aggressive power assertion, aggressive power assertion, and induction. Results of a 2 X 2 within subjects ANOVA indicated that parents were more likely to endorse the use of all types of discipline interventions when addressing aggressive violations as opposed to non-aggressive violations. Transgressions that could result in injury to self or others (safety violations) elicited higher levels of aggressive power assertion and induction than did violations of moral or social norms. Parents did not increase levels of non-aggressive discipline when addressing safety violations. The interaction between the aggressiveness and the type of the transgression was also significant, indicating that parents were more likely to use aggressive power assertion and induction for aggressive social violations than for non-aggressive safety violations.

Parental reactions to children's behaviors are moderated by a complex interaction between parent and child variables (Catron & Masters 1993; Dix, Reinhold, & Zambarano, 1990; Dix, Ruble, & Zambarano, 1989; Holden, Coleman, & Schmidt, 1995; Tisak, Nucci, & Jankowski, 1996). To understand this relationship, studies have examined parental responses to different types of child transgressions, with a variety of methods being used to classify the transgressions. While theoretical schemes for classification of transgressions are logical and meaningful, selecting items and examples to represent specific categories

---

Author info: Correspondence should be sent to: H. G. Schneider. Department of Psychology; Appalachian State University. Boone, NC 28608.
schneiderhg@appstate.edu

North American Journal of Psychology, 2002, Vol. 4. No. 2, 381-394.

© NAJP

has often been subjective rather than empirical. The purpose of the present study was twofold: first, an attempt was made to generate a set of items representing a wide range of child misbehaviors that could be used to create a useful taxonomy of transgression. The second goal of the study was to evaluate how parental disciplinary choice was influenced by the nature and severity of child transgressions.

Transgression Types

Parents determine the amount and type of discipline they use based on factors such as the age and gender of the child, level of aggression, potential for injury, parental emotional and cognitive states, and whether the offense violates social or moral standards (Catron & Masters, 1993; Holden et al., 1995). Holden et al. (1995) found that mothers were more likely to use spanking in response to aggressive behaviors compared to any other type of offense. Their classification of transgression was a simplified version of Dunn and Munn's (1987) categorization system, which included variables such as caretaking, manners and politeness, violations of the rights of others, destructive behavior, and aggressive behavior.

Tisak and Turiel (1984) used three similar categories for child transgressions: prudential violations, social convention violations, and moral violations. Prudential violations are actions that may result in physical harm to the child, such as drinking poison or playing with fire. Social convention transgressions are violations of common social rules, such as manners or dress. Finally, moral violations consist of actions that have a negative psychological or physical impact on others, such as fighting or property damage.

Catron and Masters (1993) used this system of classification to study parental use of corporal punishment and children's acceptance of discipline. Pre-school children did not differentiate between transgressions, holding a generalized view of appropriateness of parental discipline. However, fifth graders and their mothers showed more discrimination between transgressions, with both viewing prudential and moral transgressions as more deserving of all types of discipline than social transgressions. Prudential transgressions were viewed as deserving more corporal punishment than moral transgressions.

With these models of categorization, the type of transgression may be confounded by the severity of the offense. Severity, rather than the type of offense, may determine the parental reaction and disciplinary choice. To better understand the effects of transgression types, the behaviors representing these categories should have similar levels of severity or they be matched based on severity within categorization. Even if equivalent levels of severity cannot be established across the types of transgressions, then the ratings of severity may be useful as a covariate. The materials used for these studies employed behavioral descriptions or vignettes created by the researchers to logically represent categories. Empirical support for the selection of these items has not been provided; additionally, there is little evidence of inter-rated agreement about classification of the types of transgressions.

Parental responses to different transgressions vary considerably. Previous studies have mainly focused only on parental use of corporal punishment, which is one type of power assertive discipline. The literature on parental disciplinary choice neglects the study of other types of disciplinary techniques, such as induction and non-aggressive power assertion. Only Dix, et al. (1989) studied a range of parental disciplinary responses as mediated by child behavior. They focused on parental beliefs concerning children's competencies and understanding of rules, rather than the nature of the children's transgressions. Little is known about the moderating effects of children's transgression on parental use of non-aggressive discipline, even though these interventions are used far more frequently than corporal punishment (Straus & Kantor, 1994).

Parental disciplinary interventions are usually divided into two categories: power assertion and induction (Hoffman & Saltzstein, 1967; Hoffman, 1975; Eisikovits & Sagi, 1982; Rich, 1993). Induction is an intervention during which the parent explains to the child the consequences of his/her actions on self or others or the rationale behind the rule (Hoffman, 1975). The prevalence of inductive discipline use, and the factors that increase the use of induction by parents has not been well documented. Induction is contrasted to power assertive interventions. During these power assertive interventions the parent uses the power inherent within the parent-child relationship to enforce rules or to punish the child after a violation. These interventions include corporal punishment but can also less aggressive actions such as time-out and withdrawal of privileges.

Few studies have attempted to relate children's transgressions with parental discipline, and many studies concerning the effects of parental discipline style, contain problematic theoretical and methodological issues. Studies on the effect of power assertion have failed to separate aggressive from non-aggressive power assertive interventions, which is necessary since these two types of techniques may lead to significantly different outcomes, and possibly serve different purposes for the disciplinary agent. According to Hoffman (1994), the aggression component of some interventions (corporal punishment, scolding) may make the child very anxious and fearful. This elevated level of arousal may prevent some children from processing the inductive message from the parent, thereby reducing the efficacy of the intervention. Studies also suggest that aggressive interventions may be used by the parent to relieve his/her personal stress or frustration with the child, and not simply as a disciplinary intervention (Holden et al., 1995). It is clear that aggressive and non-aggressive power assertion techniques are qualitatively different, and should not be studied as a unit. Finally, parents may use a combination of discipline responses to address a specific situation. Any assessment of parental disciplinary choice must allow parents to choose different disciplinary interventions for the same violation. The present study attempted to address these theoretical and methodological issues. The first step was to develop a scaled set of items that accurately reflect current theoretical models regarding the taxonomy of children's transgressions. The next step was to assess parents' reactions to a series of child behaviors based on the types of transgressions and determine if parental disciplinary choices were influenced by the nature of the child's transgression.

Scale Development and Item Pools

Psychology graduate students and parents generated brief descriptions of children's behaviors that may result in a disciplinary response from parents. We attempted to include a very wide range of behaviors with an adequate representation of the different types of offenses with varying levels of severity. Item generation and subsequent editing resulted in an initial pool of 250 transgressions.

Participants – Sample 1. The first sample consisted of 87 college student volunteers from introductory psychology classes. Students received research credit in their classes for their participation. There were 56 females and 31 males (mean age of 18.7 years) and 92% of the subjects were white (*n*=80), 7% were black (*n*=6) and 1.1% were Hispanic (*n*=1). Participants completed a questionnaire containing the 250 behavioral transgressions presented in a set of random orders. For each transgression, the subjects were asked to rate the severity of the item on a Likert-type scale from 1 (trivial violation) to 4 (severe violation). They were also asked to assign the transgression to one of three categories: Social convention violation (e.g. lying or stealing), safety violation (e.g. playing with fire or a sharp object), and aggressive violation (e.g. hitting someone or killing an animal).

Item Selection - The 250 items were categorized based on the nature and severity of the offense. The agreement rate for the nature of the offenses was reviewed, and transgressions that were consistently assigned to one of the categories (by at least 75% of the participants) were selected as an item for that category. Within each category the severity ratings were used to select the six most and six least severe items. This selection process created six clusters of items: least-severe and most-severe for social, safety, and aggressive violations (see Table 1).

TABLE 1 Agreement Rate and Severity Ratings for Item Set No. 1

	Agreement rate	Severity rating
Safety Transgression - Low Severity		
Neglecting to use sunscreen during a sunny day at the beach	.78	1.87
Burning candles in his/her room	.81	2.09
Running around the house with a sharpened pencil	..95	2.19
Trying to catch a snake	..83	2.34
Riding his/her bike/skates without a helmet	..88	2.35
Means:	.85	2.16
Safety Transgressions - High Severity		
Playing with a gun	.87	3.81
Playing with a dangerous chemical	.96	3.66
Taking prescription medications with the intention of getting high	.85	3.63
Having unprotected sexual intercourse with boyfriend/girlfriend	.80	3.63
Taking hallucinogenic drugs	.82	3.59
Means:	.86	3.66
Aggressive Transgressions - Low Severity		
Breaking a window in anger	.81	2.52
Violently pushing a child away from his toys	.75	2.79
Getting in a fist fight with his/her sibling	.77	3.00
Kicking the family pet	.86	3.16
Beating up a younger child	.82	3.26
Means:	.80	2.94
Aggressive Transgressions - High Severity		
Pointing a gun at another person	.75	3.85
Cutting a child with a knife during a fight	.92	3.85
Hitting a teacher	.95	3.81
Beating a friend with a bat over the head	.87	3.81
Killing the family pet	.93	3.79
Means:	.88	3.82
Social Transgressions - Low severity		
Leaving the cap off the toothpaste	.98	1.12
Leaving the ice cream out	.98	1.20
Leaving the refrigerator's door open	.97	1.25
Swapping lunches at school	.92	1.27
Talking to friends on the phone longer than allowed	.98	1.31
Means:	.95	1.23
Social Violations - High Severity		
Masturbating with another child of the same age	.77	3.24
Stealing a copy of an upcoming test	.81	3.08
Trying to sell your belongings without permission	.82	3.04
Putting down someone else's culture	.88	2.96
Calling 1-900 numbers	.82	2.92
Means:	.82	3.04

The overall agreement rate for these items suggested that subjects were basically able to accurately assign the behaviors to discrete categories. For safety and social transgressions, a large number of items were endorsed, allowing for the creation of categories that clearly varied in terms of severity. However, the differences between the high and low severity ratings for aggressive violations were not significant. The number of items that were endorsed as aggressive were minimal (only 14 out of 250), which limited variance of the severity ratings.

Clear agreement that a transgression was aggressive was evident for only those items that represented extreme levels of aggression, with items that contained less aggressive actions being often classified as safety or social violations. Rather than considering aggression as a separate type of transgression, this construct may be an important dimension of all types of violations.

Subjects - Sample 2. A second study was conducted to create a classification system that included aggressiveness as a component of both social or safety violations. The same item pool of 250 transgressions was presented to a second set of college students (*N* = 48). Participants were asked to assign the transgression to one of two categories: Social violation (violations of basic social rules or norms) and safety violation (actions that might result in injury to self or others). The participants were also asked to rate the level of aggression of each item on a Likert scale from 1 (non-aggressive violation) to 4 (aggressive violation), and the severity of the item on a scale ranging from 1 (trivial violation) to 4 (severe violation).

Item Selection

Items were divided into two categories representing safety and social violations. There was clearer agreement among raters regarding this classification. Only those items with an agreement rating of .90 or higher were selected for each category. After initial analysis, it was determined that ratings of aggression and severity were very highly correlated (r = .90), which precluded using both aggressive and severity ratings as a basis for categorization. Items were selected, therefore, based on levels of aggression, since the aggressiveness of transgressions was assumed to be centrally related to disciplinary choice. The 10 least and 10 most aggressive items were included, creating four clusters of items: non-aggressive safety violations, aggressive safety violations, non-aggressive social violations, and aggressive social violations.

This group of 40 items was then evaluated based on content. Five items per cluster were selected, which represented violations applicable to most social classes and geographical locations. The final pool contained a total of 20 items, which are presented in Table 2: these items were used for the Parent Discipline Scale. This scale included parental responses on the use of induction, aggressive power assertion, and non-aggressive power assertion for each of the child transgressions. For each class of disciplinary response, two specific parental responses were generated reflecting common disciplinary actions.

TABLE 2 Agreement Rate, and Severity and Aggression Ratings for Item Set No. 2

	Agreement rate	Severity rating (*M* = 2.79) (*SD* = .85)	Aggression rating (*M* = 2.54) (*SD* = .94)
Safety Non-aggressive Transgression			
Crossing the road without looking	..90	3.17	2.12
Climbing on a tree branch that is about to break	..90	3.05	2.15
Riding his/her bike/skates without his/her helmet	..90	2.5	2.07
Trying to get something of the oven without gloves	..95	2.85	1.97
Running barefoot w/ several sharp objects on the floor	. 93	2.82	2.25
Means:	..92	2.88	2.11
Safety Aggressive Transgressions			
Cutting a child with a knife during a fight	..90	3.8	3.80
Pointing a gun at someone	..93	3.8	3.77
Beating up a younger child	..93	3.5	3.70
Severely shaking a baby	..93	3.7	3.62
Aiming and shooting fireworks at other people	..90	3.67	3.60
Means:	..918	3.69	3.70
Social Non-aggressive Transgressions			
Getting up late and missing the school bus	..97	1.7	1.32
Eating with his/her fingers	..97	1.27	1.17
Leaving the cap off the toothpaste	..90	1.27	1.27
Watching TV when he/she should have been doing homework	1.0	1.67	1.37
Leaving the refrigerator's door open	..97	1.62	1.40
Means:	..96	1.51	1.31
Social Aggressive Transgressions			
Vandalizing school property with his peers	..90	3.55	3.25
Breaking his/her friend's toy on purpose	..97	2.80	3.20
Using profanity towards a teacher	..9	3.32	3.10
Picking on a younger child until the child cries	.90	2.95	2.95
Intentionally scratching new furniture	1.0	2.80	2.85
Means:	.95	3.08	3.07

THE PARENT DISCIPLINE SCALE – APPLICATION

Participants

One hundred and one parents participated in the study. Sixty-seven percent were mothers (*n*=68), and 33% were fathers (*n*=33). The sample consisted of mostly middle class families with an average education of 15 years, and an average of 2.5 children. Ninety-four percent were white (*n*=97), 1.9% were African American (*n*=2), 1% were Hispanic (*n*=1), and 1% were Asian (*n*=1).

Design and Materials

Parents of college students were contacted by mail and were asked to complete the PDS. If they were willing to participate, they signed a consent form, and completed the testing packet. The 20 behavioral transgressions were presented to parents in a random order, and parents were asked to rate how likely they would be to use six specific disciplining interventions if their child had engaged in these behaviors. They were asked to assume that they were making these ratings when their child was between 8 and 13 years of age. The forced choice nature of the questions may lead either to more socially desirable responses or to a greater consistency than might be seen in actual behavior. The six specific interventions represented three types of discipline: Aggressive power assertion (use of corporal punishment, severe scolding), non-aggressive power assertion (withdrawal of privileges, timeout or grounding), and induction (explaining the consequences that his/her behavior have on self and others). Parents rated their predicted responses on a 5-point Likert-type scale. The PDS produced separate scores on levels of aggressive power assertion, non-aggressive power assertion, and induction. Using the pre-determined classification of items, the scale also allows for analysis of the amount of each discipline style used for each type of transgression. Appendix A contains the full scale as it was presented to parents.

RESULTS

Two paired sample T-tests were conducted to determine if there were significant differences between the three types of discipline styles. The mean endorsement score for overall use of induction (M = 162.21) was significantly higher than for non-aggressive power assertion (M = 113.81), $t_{(100)} = -19.00$, $p < .001$. Use of induction was also higher than for the use of aggressive power assertion (M = 77.66), $t_{(100)} = -12.68$, $p < .001$. Table 3 presents the mean endorsement scores for each discipline style during all types of transgressions. Parents reported most frequent use of induction, followed by non-aggressive discipline, with aggressive discipline being reported least frequently.

TABLE 3 Means of all Discipline Interventions for each Type of Transgression.

	Aggressive Power assertion		Non-aggressive Power assertion		Induction	
Transgression	*M*	*SD*	*M*	*SD*	*M*	*SD*
Safety						
Aggressive	26.94	11.17	38.23	9.49	47.45	4.85
Non-aggressive	14.81	5.32	18.74	8.07	38.38	7.24
Social						
Aggressive	22.52	9.38	37.06	9.02	44.83	5.96
Non-aggressive	13.39	4.20	19.86	7.19	31.55	9.43

A 2 x 2 within subjects factorial analysis was conducted to evaluate the effects of different types of transgressions on parental choice of disciplinary intervention. The within subjects variables were type of transgression (social vs. safety), and aggressiveness of transgression (aggressive vs. non-aggressive). The dependent measures were the likelihood of the use of each type of discipline (aggressive power assertion, non-aggressive power assertion, and induction). Main and interaction effects were tested using the multivariate criterion of Wilks' Lambda (Λ).

The main effect of type of transgression was significant for aggressive power assertion ($\Lambda = .59$, $F_{(1,100)} = 69.97$, $p < .001$), and induction ($\Lambda = .51$, $F_{(1,100)} = 95.58$, $p < .001$). This effect was not significant for non-aggressive power assertion ($\Lambda = 1.00$, $F_{(1,100)} = .007$, $p = .935$). Parents reported that they were more likely to use aggressive power assertion and induction to address safety violations than social violations. However, parents predicted they would use the same levels of non-aggressive power assertion to address both safety and social violations.

The main effect for aggressiveness of transgression was significant for aggressive power assertion $\Lambda = .35$, $F_{(1,100)} = 188.07$, $p < .001$, non-aggressive power assertion $\Lambda = .16$, $F_{(1,100)} = 510.77$, $p < .001$, and induction $\Lambda = .27$, $F_{(1,100)} = 265.06$, $p < .001$. Parents reported that they were more likely to use all types of discipline to address aggressive violations rather than non-aggressive violations.

The interaction between type of transgression and discipline type was significant for aggressive power assertion $\Lambda = .84$, $F_{(1,100)} = 18.67$, $p < .001$, non-aggressive power assertion $\Lambda = .90$, $F_{(1,100)} = 10.67$, $p = .001$, and induction $\Lambda = .81$, $F_{(1,100)} = 22.84$, $p < .001$. Paired sample *t*-tests indicated that the effect of type of transgression sometimes was mediated by the degree of aggressiveness of the transgression. Although parents endorsed using more aggressive power assertion and induction to address safety violations over social transgressions, this pattern changed when the social violation was aggressive. Parents endorsed using more induction and aggressive power assertion to address *aggressive social* violations than *non-aggressive safety* violations $t_{(100)} = -24.35$, $p < .001$; $t_{(100)} = -10.18$, $p < .001$. After controlling for family wise error this effect did not apply to the use of non-aggressive power assertion.

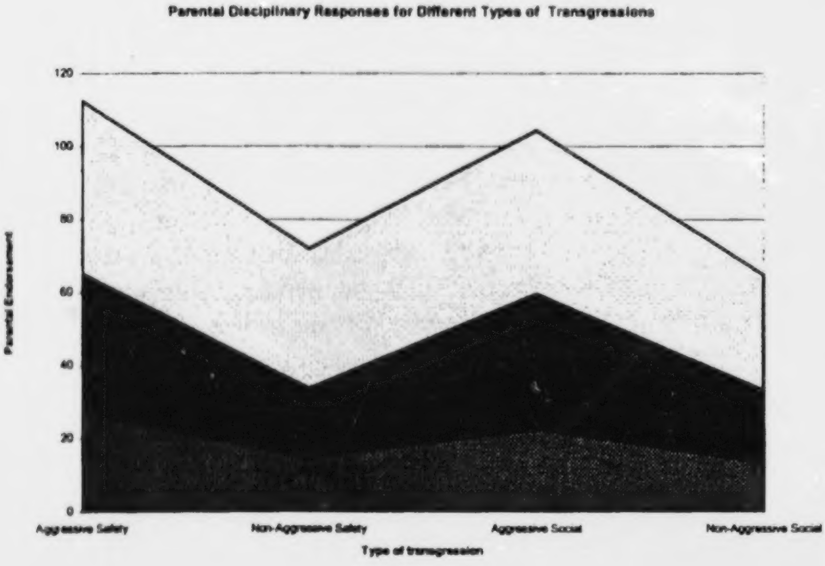

FIGURE 1 Parental Discipline Responses by Transgression Type

DISCUSSION

A taxonomy of behavioral transgressions was created based on the nature of the violation (safety or social norm violations) and the aggressiveness of the child's action. This combined classification was developed because the initial classification scheme (safety, social conventions or aggression) was confounded by severity. For safety and social transgressions, it was possible to create item sets differing in severity, but most of the aggressive transgressions were viewed as being severe. Aggression appears to be the most important dimension, since both severity ratings and parental disciplinary responses varied based on this dimension.

BEST COPY AVAILABLE

Parents in this study reported that the type of discipline intervention chosen, vary according to whether the child's transgression was a social or safety violation. Parents endorsed using more induction and aggressive power assertion as reaction to safety violations than social violations. Catron and Masters (1993) reported similar results suggesting that parents were more likely to use aggressive power assertion techniques and induction when addressing safety compared to social violations. However, parents in the current study did not vary their use of non-aggressive power assertion as a function of type of transgression. This indicates that parents either find little value in increasing the rate of non-aggressive techniques to address safety violations, or these disciplinary interventions serve different purposes for the parent.

The perceived danger inherent in safety violations may account for the increased use of induction. By using induction, the parent ensures that the child understands the consequences of his/her actions. Similarly, the increased use of aggressive power assertion could be attributed to a desire to use more powerful discipline in response to safety violations. However, these transgressions may also elicit a strong negative emotional reaction on the part of the parent, which can increase the possibility of responding aggressively to the child. Parents have been found to be more likely to use corporal punishment when they are in a negative emotional state than when their emotions are more regulated (Holden et al., 1995).

Parents also reported to be more likely to utilize all types of discipline when the behavioral transgressions were aggressive. The significant effects of aggressiveness of the transgression on disciplinary response are consistent with the findings of previous studies concerning the determinants of corporal punishment (Holden et al., 1995). It appears that the aggressiveness was more important than the type of transgression in determining parental selection of disciplinary intervention. Parents appeared to consider aggressive violations as requiring more discipline than safety violations. However, this was not due to parents considering aggressive violations to be more severe than all non-aggressive violations. Severity ratings indicated that parents considered aggressive social and non-aggressive safety violations to be equally severe, yet they reported using more induction and aggressive power assertion to address aggressive social violations than non-aggressive safety violations. The overall value system of the parent may be related to the selection of discipline style. Parents may provide more severe discipline to actions that violate social or religious values, such as "one should not be aggressive towards others". Adults may think that aggressive social violations required more discipline than non-aggressive safety violations because aggressive violations violate their value system, even when both types of violations are considered to be equally severe.

Although the PDS appears to provide useful information about the manner in which behavioral transgressions affect parental choice of disciplinary interventions, the data on validity and reliability are limited. The reported frequencies of use for different discipline interventions are consistent with those found in large epidemiological studies (Straus & Kantor, 1994) suggesting some construct validity. Differences found between discipline styles as mediated by transgression type are also consistent with previous findings (Catron & Masters, 1993; Holden et al., 1995). Both the PDS and the *Conflict Tactics Scale* (CTSPC; Straus, 1991) were used in a study of discipline styles and moral reasoning (Lopez, Bonenberger, & Schneider, 1999), with the aggressive power assertion scale of the PDS significantly correlated with the physical assault sub-scale of the CTSPC. In addition, the non-aggressive power assertion scale was correlated with the non-violent discipline sub-scale of the CTSPC.

Future studies may attempt to evaluate properties of safety and aggressive transgressions that elicit high levels of discipline, as well as the role of parental affect on the selection of disciplinary choice. Understanding these interactions is essential for the study of parenting styles and the effects of disciplinary interventions on social, psychological and emotional functioning.

REFERENCES

Catron, T. F. & Masters, J. C. (1993). Mothers' and children's conceptualizations of corporal punishment. *Child Development, 64, 6,* 1815-1828.

Dix, T., Reinhold, D. P., & Zambarano, R. J. (1990). Mothers' judgement in moments of anger. *Merrill-Palmer Quarterly, 36, 4,* 465-486.

Dix, T., Ruble, D. N., & Zambarano, R. J. (1989). Mothers' implicit theories of discipline: Child effects, parent effects, and the attribution process. *Child Development, 60, 6,* 1373-1391.

Dunn, J. & Munn, P. (1987). Development of justification in disputes with mother and sibling. *Developmental Psychology, 23, 6,* 791-798.

Eisikovits, Z. & Sagi, A. (1982). Moral development and discipline encounter in delinquent and nondelinquent adolescents. *Journal of Youth and Adolescence, 11, 3,* 217-230.

Hoffman, M. L. (1975). Moral internalization, parental power, and the nature of parent-child interaction. *Developmental Psychology, 11, 2,* 228-239.

Hoffman, M. L. (1994). Discipline and internalization. *Developmental Psychology, 30, 1,* 26-28.

Hoffman, M. L. & Saltzstein, H. D. (1967). Parent discipline and the child's moral development. *Journal of Personality and Social Psychology, 5, 1*, 45-57.

Holden, G. W., Coleman, S. M. & Schmidt, K, L. (1995). Why 3-year-old children get spanked: Parent and child determinants as reported by college-educated mothers. *Merill-Palmer Quarterly, 41, 4*, 431-452.

Lopez, N. L., Bonenberger, J., & Schneider, H. G. (1999). Parental discipline styles and the development of empathy and moral thought. Presented at the 2nd International Conference: The Changing Family and Child Development. Banff, Canada.

Rich, J. M. (1993). Discipline and moral development. *The High School Journal, 76, 2*, 139-144.

Straus. M. A. (1991). Discipline and deviance: Physical punishment of children and violence and other crime in adulthood. *Social Problems, 38, 2*, 133-154.

Straus, M. A. & Kantor, G. K. (1994). Corporal punishment of adolescents by parents: A risk factor in the epidemiology of depression, suicide, alcohol abuse, child abuse, and wife beating. *Adolescence, 29, 115*, 543-561.

Tisak, M. S., & Turiel. E. (1984). Children's conceptions of moral and prudential rules. *Child Development, 55, 3*, 1030-1039.

Tisak, M. S., Nucci, L. P., & Jankowski, A. M. (1996). Preschool children's social interactions involving moral and prudential transgressions: An observational study. *Early Education and Development, 7, 2*, 137-148.

Requests for copies of the Parent Discipline Scale should be addressed to H. G. Schneider. Department of Psychology; Appalachian State University. Boone, NC 28608. schneiderhg@appstate.edu

The Effect of Autobiographical Writing on the Subjective Well-Being of Older Adults

Nancy Richeson
University of Southern Maine

James A. Thorson
University of Nebraska at Omaha

Our purpose was to examine the effects of autobiographical writing on the well-being of older adults to determine if there were differences between participants who enrolled in an autobiographical writing class or those who took part in a continuing education liberal arts class. The *Positive and Negative Affect Scale* was used, as was the *Satisfaction with Life Scale*. It was hypothesized that there would be a significant difference between those who took part in the autobiographical writing classes and those who participated in the liberal arts classes on three dependent measures of subjective well-being. A split-plot analysis of variance was used, and the results indicated a significant difference between the scores on negative affect achieved by older persons who participated in the autobiographical writing classes when compared to those in the liberal arts classes. The negative affect of those enrolled in the autobiographical writing classes decreased from pretest to posttest when compared to the liberal arts class participants. No significant difference was found from pretest to posttest on the other two dependent measures, positive affect and satisfaction with life.

> "Know well, Socrates," Cephalus said, "that when a man faces the thought that he must die, he feels fear and anxiety about that which did not trouble him before." Plato's *Republic*, I 330d

The theory of the life review (Butler, 1963) is one that has stood the test of time in the field of gerontology. Briefly, life reviewing involves a return to consciousness of past life events among persons entering later life. Brought on by a realization of impending death, the life review is characterized by spontaneous reminiscing and a quest to put life events into context. Righting old wrongs, paying debts, and assessing one's past deeds are activities the old merchant Cephalus spoke of in his dialogue with Socrates as being brought on in late life along with a realization of one's own finitude. Life reviewing may be part of a search for meaning in old age.

Author info: Correspondence should be sent to Nancy Richeson, University of Southern Maine, G-26 Masteron Hall, 96 Falmouth Street, Portland, Maine 04104-9300; email to: richeson@usm.maine.edu

North American Journal of Psychology, 2002, Vol. 4, No. 3, 395-404.

© NAJP

Theoretically, the process of the life review is therapeutic: most people are able to rationalize past conflicts and come to the conclusion that they've led pretty good lives. The serenity often associated with old age is thought to be a product of positive life reviewing. Because of this acceptance of the beneficial outcomes of life reviewing, reminiscence therapy has become popular in recent years. Karel and Hinrichsen (2000) present empirical evidence of positive outcomes for life review therapy. Many programs for seniors have implemented programs providing structured life-reviewing projects, including autobiographical writing classes emphasizing reminiscence.

The present study examines the effects of an autobiographical writing class on the subjective well-being of older adults enrolled in Institutes for Learning in Retirement, an affiliate program of the Elderhostel Network in the United States and Canada. The design of the autobiographical writing classes is based on the work of Birren and Deutchman (1991), and is a semi-structured approach to life reviewing.

Life reviewing, says Robert N. Butler (1963), is common, normal, and necessary. It is a universal occurrence in later life, part of the aging process. It involves a conscious return to past experiences, reviewing unresolved conflicts and integrating them into a deeper understanding of oneself and the meaning of one's life. It has been suggested that it in a way is part of the process of coming to terms with one's creator.

Butler, Lewis, and Sunderland (1998) noted that the process of the life review leads to self-awareness and self-acceptance, and hence a more meaningful view of one's life. This understanding prepares one for the end of life by decreasing one's fears and anxieties about death (Thorson & Powell, 1990). The life review can help to transform the way individuals view their lives, gaining a more positive view of self and greater life satisfaction. The life review is an examination of both positive and negative experiences, and the goal of life reviewing is to gain a better understanding and appreciation of one's life (Havighurst & Glasser, 1972; Romaniuk & Romaniuk, 1981).

More recent research has supported the use of life reviewing as a cost-effective intervention for the treatment of depression for older adults in nursing homes (Jones & Beck-Little, 2002). Furthermore, Haight, Michel, & Hendrix (2000) have reported a continued and significant improvement over time for those participants in a nursing home who received a life review intervention on measures of depression, life satisfaction, and self-esteem. This suggests that the life review intervention is effective in improving the well-being of older adults.

Autobiography, or the written story of a person's life, can consist of short writings that focus on one event or memory in an individual's life (Atkinson, 1995). Guided autobiography, as defined by Birren and

Deutchman (1991), is the process of reviewing one's life using provocative themes and questions that promote reflective thinking. A group leader who sequentially presents the themes in order to gently guide participants into their memories and experiences facilitates the process. According to Birren and Deutchman, the autobiographical writing process can deviate from a formal format and still be beneficial to older adults. Their method can be adapted to groups as well as individuals, making it ideal for use with older adults enrolled in continuing education programs.

Only a few studies have examined autobiographical writing groups as an intervention for enhancing well-being, and their findings are mixed. Malde (1988) examined the effects of an autobiographical writing intervention using two treatment groups and one control group — a total of 39 participants. After five weeks there were no significant changes on measures of well-being (*Tennessee Self-concept, Time Competence Scale, Personal Orientation Inventory,* and the *Purpose of Life Test*), using a post-test only research design.

However, a follow-up study was conduced one year after the course was completed that surveyed the 39 participants. Thirty-three (76%) of the questionnaires were returned. Although the original measures were not used the questionnaire that was developed tapped into similar concepts. The questionnaire included five items that used a 7-point scale and measured self-concept, time competence, and purpose in life. This study demonstrated that autobiographical writing did promote a statistically significant positive change in self-concept, time competence, and purpose of one's life. Malde concluded that the effects of this type of writing promote changes in well-being, but the changes may not be immediate, and participants may need additional time to integrate what they have observed and experienced.

Barry (1988) examined the effect of autobiographical writing groups on the oldest of the old (i.e., 85 years and older), who are often left out of activities in senior centers. The intervention followed the format developed by Birren and Deutchman. Barry found that the most important outcome of the group was an increase in self-esteem; depression was also reduced, and self-acceptance increased. Another outcome of the treatment was strengthened family connections as stories were shared.

Botella and Feizas (1993) reported that an autobiographical writing class for eight participants over a three-month period promoted the reconstruction of the older adults' life experiences by enhancing self-awareness and encouraging self-disclosure. The class members showed significant improvements in self-esteem and noted that they had learned to view life experiences in new ways.

Thus, a number of articles in the literature suggested an association between structured life reviewing through autobiographical writing and various measure of self-esteem. A more comprehensive study of these concepts seemed to be appropriate.

METHOD

The present study used a quasi-experimental design with a nonequivalent control group to examine the effects of autobiographical writing on older adults' well-being. The participants were persons 50 to 85 years of age who enrolled in classes in autobiographical writing (the experimental group) and those enrolled in liberal arts classes (the control group) through the Institutes for Learning in Retirement (ILR) presented by Elderhostel programs in the United States and Canada. A total of 16 autobiographical writing classes and nine liberal arts classes were selected to participate in the study. Autobiographical writing classes were defined as those classes that ask the participants to write and share with the class a selection of short autobiographical life stories. Liberal arts classes had different content that did not have a autobiographical, creative, or narrative writing component. Approval for the study was obtained from the University of Nebraska Institutional Review Board for the Protection of Human Subjects.

The course facilitators were recruited by posting a message on an e-mail discussion group for members and administrators of the Elderhostel-affiliated Institutes for Learning in Retirement (Rodgers, 1999). An explanation of the study's goals was provided, and volunteers who responded were screened and enrolled if they met the criteria. An additional mailing requesting participation was sent to all ILRs. Course facilitators were enrolled into the experimental group if their class: 1) met for two hours once a week for six to ten weeks, 2) was assigned at least a two-page written paper, revolving around a particular theme, 3) had the opportunity to share their written work with the class, and 4) was facilitated by a peer leader. Participants were enrolled into the control group if: 1) they had registered for a liberal arts class that met for 2 hours once a week for 6 to 10 weeks, 2) they were not currently enrolled in any type of writing course, and 3) they had no prior experience with autobiographical writing classes. Participants were living independently in their home communities.

A packet of instructions, data collection materials, a letter from the researcher, and a scripted description of the study was sent to the course instructors who had agreed to distribute and collect the information. To guarantee the integrity of the data collection process prior to the beginning and toward the end of the study, e-mails and phone calls were made to ensure that the facilitators were clear on how to distribute and collect the data.

On the first day of each class the course facilitator presented the members in the classes with a presentation that included a letter from the researcher, a scripted description of the study, an informed consent form, and the pretest (i.e., demographic information and outcome measurement tools). The individual members of each course made a decision on the first day of class if they would like participate in the study. If the members of the class decided to participate, the pretest data was distributed. On the last day of class the participants completed the posttest data. All materials were collected by the course facilitators and mailed to the researcher. All material was coded to protect the confidentiality of the participants. A total of 400 older adults agreed to participate.

Individuals could not be randomly assigned to groups because each chose to register for their own courses.

Instruments

Subjective well-being was measured using the *Positive and Negative Affect Scale* (PANAS) and the *Satisfaction with Life Scale* (SWLS). The PANAS (Watson, Clark, & Tellegen, 1988) is a 20-item psychometric scale that assesses well-being. Items are scored on a Likert scale ranging from very slightly or not at all (1) to extremely (5). The scale consists of a number of words that describe various feelings and emotions; respondents are asked to indicate the extent to which they felt that way during the past few days. The PANAS has shown good reliability. Cronbach alphas were .88 for its Positive Affect scale and .95 for its Negative Affect scale in the Watson et al. study.

The *Satisfaction with Life Scale* (Diener, Emmons, Larsen, & Griffin, 1985) is used to measure global life satisfaction, a component of subjective well-being. The SWLS has five questions. Respondents are asked to read each statement and indicate their agreement with it on a seven-point continuium ranging from strongly disagree (1) to strongly agree (7). Participants were asked to be open and honest in their responses (Shevlin, Brunsden, & Miles, 1997). The SWLS has shown high internal consistency and high test-retest reliability. Scores on the SWLS correlated with other measures of subjective well-being, and correlated predictably with specific personality characteristics (Diener et al). Its Cronbach alpha was .87 in the present study.

The demographic data collected included: age, marital status, years of education, overall health status, gender, ethnicity, and yearly income. The experimental group was asked how many times they had taken an autobiographical writing class.

RESULTS

A confidence level of .05 was used for this study. A total of 374 of the initial respondents completed both the pretest and the posttest. There were 150 participants in the autobiographical writing classes and 224 controls in the liberal arts courses. Those in the autobiographical writing classes were somewhat older (M = 70.4 yrs. vs. 68.7 yrs.). A total of 92% of both groups were Caucasian. A total of 76% of the participants in autobiographical writing classes and 55% of the liberal arts class participants reported that they were currently married. The majority of participants, 48% stated their health was good, with 38% indicating that they were in excellent health. On average, the two groups had 16 years of education.

TABLE 1 Mean and Standard Error Scores for Pretest and Posttests of Positive Affect, Negative Affect, and Satisfaction with Life

Variables	M	SE
Positive Affect		
AW Classes		
Pretest	35.846	.548
Posttest	35.333	.586
LA Classes		
Pretest	35.583	.608
Posttest	35.365	.649
Negative Affect		
AW Classes		
Pretest	16.659	.509
Posttest	14.671	.501
LA Classes		
Pretest	14.671	.460
Posttest	15.399	.453
Satisfaction with Life Scale		
AW Classes		
Pretest	24.493	.529
Posttest	25.260	.526
LA Classes		
Pretest	25.260	.470
Posttest	25.492	.467

Table 1 presents the mean pretest and posttest scores on the *Positive Affect, Negative Affect,* and *Satisfaction with Life* scales for the participants in the two groups. There were nonsignificant declines in Positive Affect for both groups' pre and post-tests and nonsignificant increases in life satisfaction for both groups. There was a more meaningful decline in Negative Affect for the autobiographical writing participants, and an increase in Negative Affect for the liberal arts participants.

TABLE 2 Split-plot Analysis of Variance for Negative Affect

Variables	DF	F	SS	MS	Sig.	Eta Sq.
Between Subjects NA	1	2.476	85.369	85.369	.117	.009
Error	285		9825.31	34.475		
Within Subjects Factor (Pre- Post)	1	1.065	33.398	33.398	.303	.023
Error	285		8935.25	31.352		
Within Subjects Interaction	1	6.663	208.903	208.903	.010*	.023
Error	285		8935.25	31.352		
Total NA	2		327.67			

*$p < .05$

Table 2 presents the split-plot analysis of variance for the between-subjects and within-subjects factors, which indicates that the decline in scores on Negative Affect for the autobiographical writing group was significant.

Table 3 presents the post-hoc test, Tukey's HSD, and shows a significantly greater change in the Negative Affect of older adults who participated in the autobiographical writing classes than participants in the liberal arts courses.

TABLE 3 Summary of Tukey HSD Test of Analysis

Contrast	*DF*	*M* Differences	*MS*	*Q*
Liberal Arts Class	285	1.683784	31.352	1.560

Autobiographical Writing Class	285	1.683784	31.352	3.63*

*$p < .05$

DISCUSSION

In this study, autobiographical writing classes, a structured approach to life reviewing, did not significantly alter *Positive Affect* or *Satisfaction with Life* scores of the older adult respondents. However, participation in the autobiographical writing course was associated with a decrease in respondents' scores on *Negative Affect*.

While the sample of the study was a large one (n = 374), the quasi-experimental nature of the design posed some internal validity problems. For example, causal ambiguity was a problem. Even though the course facilitators were screened to determine if their classes met certain criteria, there was no way of determining the quality of the teaching. One facilitator, say, may have been more skilled at creating a supportive environment that encouraged an open exchange of ideas and feelings, enabling a more thoughtful exploration of one's life. Also, there was a self-selection bias. The participants enrolled in either the autobiographical writing or liberal arts courses because they had an interest or expertise in the subject. Other extraneous variables that could have influenced the outcomes of the study are the conditions of the classrooms, the interactions the participants had with the other members of the class, and encouragement of the participants family members.

Additionally, the life-satisfaction and positive affect of the participants in this study may not have been influenced by the intervention more than they were influenced by the demographic characteristics of the participants. For the most part they reported they were married, in good to excellent health, had higher than average income, and were well educated. All these variables could contribute to the well-being of older-adults (Diener, et al, 1999). Malde (1988), reported similar demographic characteristics with inconclusive findings in her initial research project. Previous research in life reviewing has mostly been completed on vulnerable older adult populations who may be more susceptible to depressive symptoms (Jones & Beck-Little, 2002).

There was a significant decrease in negative affect among the participants in the autobiographical writing classes. Previous studies of autobiographical writing interventions have not addressed negative affect. However, Haight, et al. (2000), in their study of nursing home residents, found a decrease in depression using the *Beck Depression Inventory* as the measurement tool. Screening measures that monitor negative affect have previously been overlooked in research pertaining to autobiographical writing using an education-orientated approach.

As the number of older adults in the population continues to grow and Elderhostel programs expand, there will be a sizeable number of older people seeking for educational outlets that provide specified outcomes, perhaps a greater sense of well-being. Continuing to study which outcomes tend to produce intended results might be seen as an important part of the educational process. Given the interest in education for older adults, and specifically autobiographical writing classes, more research in the area is warranted, exploring the issues of reminiscence and life reviewing and their possible decrease in negative affect among the elderly.

REFERENCES

Atkinson, R. (1995). *The gift of stories: Practical and spiritual application of autobiography, life stories, and personal mythmaking*. Westport, Connecticut: Bergin and Garvey.

Barry, J. (1988). Autobiographical writing: An effective tool for practices with the oldest old. *Social Work, 33*, 449-451.

Birren, J. E., & Deutchman, D. E. (1991). *Guiding autobiography groups for older adults: Exploring the fabric of life*. Baltimore: Johns Hopkins University Press.

Botella, L., & Feizas, G. (1993). The autobiographical group: A tool for the reconstruction of past life experience with the aged. *International Journal of Aging and Human Development, 36*, 303-319.

Butler, R. N. (1963). The life review: An interpretation of reminiscence in the aged. *Psychiatry, 26*, 486-496.

Butler, R. N., Lewis, M. R., & Sunderland, T. (1998). *Aging and mental health: Positive psychosocial approaches* (5th Ed.). Boston: Allyn & Bacon.

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment, 49*, 71-75.

Diener, E., Suh, E.M., Lucas, R., Smith, H. (1999). Subjective well-being: Three decades of progress. *Psychological Bulletin,125*(2), 276-302.

Haight, B., Michel, Y, Hendrix, S. (2000). The extended effects of the life review in nursing home residents. *International Journal of Aging and Human Development, 50*(2), 151-168.

Havighurst, R. J., & Glasser, R. (1972). An exploratory study of reminiscence. *Journal of Gerontology, 27*, 245-253.

Jones, E.D., & Beck-Little, R. (2002). The use of reminiscence therapy for the treatment of depression in rural-dwelling older adults. *Issues of Mental Health Nursing, 23*, 279-290.

Karel, M. J., & Hinrichsen, G. (2000). Treatment of depression in late life: Psychotherapeutic interventions. *Clinical Psychology Review, 20*, 707-729.

Malde, S. (1988). Guided autobiography: A counseling tool for older adults. *Journal of Counseling and Development, 66*, 290-293.

Rodgers, D. (1999). *Elderhostel Institute Network Directory*. Boston: Elderhostel, Inc.

Romaniuk, M., & Romaniuk, J. (1981). Looking back: An analysis of reminiscence functions and triggers. *Experimental Aging Research, 7*, 477-489.

Shevlin, M., Brunsden, J., & Miles, J. (1997). Satisfaction with life scale: Analysis of factorial invariance, mean structures, and reliability. *Personality and Individual Differences, 25*, 911-916.

Thorson, J. A., & Powell, F. C. (1990). Meanings of death and intrinsic religiosity. *Journal of Clinical Psychology, 46*, 379-391.

Watson, D., Clark, L., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54*, 1063-1070.

Psychology Training Regarding HIV/AIDS Revisited

Carlos A. Escoto
Mississippi State University

Two hundred-two graduate programs in psychology were surveyed to assess the degree of progress made in HIV/AIDS training since a 1987 survey. Results indicate that in general HIV/AIDS training is now present in approximately one half of all programs surveyed and that AIDS prevention course coverage has increased. However, 31% of the programs surveyed offered no training regarding HIV/AIDS. This is a substantial improvement over the 75% of graduate programs that offered no training regarding HIV/AIDS in 1987. There are several other training issues to address. While there are a greater number of students currently involved in HIV/AIDS research, the number of faculty involved in HIV/AIDS-related research and clinical activities has not increased. The dearth of faculty members involved in training opportunities is of continued concern. Active recruitment of faculty interested in HIV/AIDS research and teaching is necessary for continued training and service enhancement for HIV/AIDS treatment and/or HIV prevention.

As Ferrando (1998) states " it is incumbent on those of us in the behavioral sciences to continue efforts toward keeping pace and integrating more fully with accelerating medical advances." Research and training are inextricably tied. Without basic research, training of emerging psychologists will be unaffected and the net amount of research will remain static. It is reasonable to believe that a lack of training at the graduate school level is impacting the amount of research being conducted on HIV/AIDS. Research has shown that training regarding HIV and AIDS in psychology, social work, counseling and psychiatry programs are lacking.

Campos, Brasfield and Kelly (1987) conducted a survey of APA accredited graduate programs regarding HIV and AIDS training in 1987. The survey showed that 75% of the responding graduate programs' curricula did not cover HIV/AIDS. Further, the authors found that other key areas where HIV/AIDS information could be provided were also deficient. For example, 51% of the responding programs did not offer any courses on human sexuality.

Author info: Correspondence should be sent to: Dr. Carlos Escoto, Dept. of Psychology, Miss. State U., Mississippi State, MS 39762.

North American Journal of Psychology, 2002, Vol. 4, No. 3, 405-416.
© NAJP

Pingitore and Morrison (1993) reassessed AIDS-related academic and clinical activities in 1989. The authors concluded that psychology programs were engaged in many AIDS-related activities. Furthermore, the authors stated that psychology departments were successful in establishing cooperative relationships with hospitals for AIDS-related training activities. It should be noted however that this study's findings are not directly comparable with the study conducted by Campos, Brasfield and Kelly (1989). Pingitore and Morrison used a different survey form than the Campos et al.(1989) study. Also, the information gathered was divergent enough from the original study to warrant an additional investigation. More importantly, this study assessed change only two years after the Campos et al. study (1989).

In 1997, the American Psychological Association (APA) surveyed graduate teaching faculty who offer HIV-related courses in various areas of psychology (Anderson, J., Campos, P. & Hamid, G., 1998). As with the previous study (Campos et al., 1989), only a small percentage (14%) of courses had HIV/AIDS as their primary focus. HIV/AIDS information was provided by approximately 54% of the respondents as part of other psychology courses. More recently, a content analysis of HIV and AIDS information in psychology textbooks, was conducted by Wong, Duffy, Faulring, Eggleston and Harper (2001). The authors concluded that while HIV information could be found in most of the texts reviewed; the information provided did not address the contributions psychology could make to curb the HIV/AIDS pandemic.

Similar findings have been reported in other clinical fields. McDaniel, Purcell, Farber, Bondurant, Donovan, Chang, & Campos (1998) reported results of a survey on HIV/AIDS knowledge in psychiatrists. The survey found deficiencies in the knowledge of psychiatrists in regards to the neuropsychological manifestations of HIV/AIDS, working with patients who are dealing with death and normal sexuality. Diaz & Kelly (1991) surveyed social work programs using the same instrument used by Campos et al. (1989). The authors reported that 65% of the responding programs provided HIV/AIDS training as part of another course or colloquia. Likewise, Hunt (1996) assessed counseling education programs. Fifty-one percent reported using colloquia to provide HIV/AIDS information. What should be noted is that previous research has focused on clinical and counseling graduate programs; experimental graduate programs also are responsible for conducting research that encompasses the area of HIV and AIDS and have not been assessed.

Evidence for the effect of training in psychology can be found in the amount of research being conducted on HIV and AIDS. "AIDS in the World" (Morton, 1992) reported that a Medline search showed that 10%

of all research published since the beginning of the HIV pandemic focuses on the psychological impact of HIV/AIDS. Ferrando (1997) reports that at a conference on retroviruses, out of 1,000 presentations and abstracts only 20 focused on the behavioral/psychological aspects of HIV/AIDS. There appears to be a rift between medicine and the behavioral or social issues related to HIV/AIDS.

A special committee of APA recommended that psychologists take a more active role in the prevention and treatment of HIV/AIDS. The APA committee estimated that research in the area of behavioral modification could cut the number of new HIV infections in half (Michaelson, 1993). Kelly, Murphy, Sikkema and Kalichman (1993) noted that because behavior change is the only method for prevention of new HIV infections, psychology should take a leading role in prevention efforts. However, the response of psychologists has not been proportionate to the severity of this health crisis. Given the lack of a successful preventative vaccine, psychology's contribution to the AIDS pandemic could be substantial.

Another area where HIV/AIDS research is needed is on the psychological effects of a diagnosis of HIV or AIDS. As Kelly and Murphy (1992) point out, it will prove to be essential to identify the different coping needs of the varied groups of persons effected by HIV/AIDS; so as to develop specific interventions to help persons infected with HIV. Developing gender and culturally specific psychological interventions is needed to address the pressing needs of persons with HIV/AIDS. The emerging field of Psychoneuroimmunology points to the ability of psychological interventions to impact immunological function which could alter disease course in persons with HIV and AIDS.(Cohen & Herbert, 1996).

Great strides have been made in the medical treatment of HIV and AIDS. Drug cocktails have resulted in a decrease in mortality from AIDS as well as a decrease in the number of persons who progress from HIV to AIDS (CDC, 1998). However, researchers suggest that psychologists focus on the development of strategies to promote adherence to the potentially complicated HIV/AIDS drug regimens (Kelly, Otto-Salaj, Sikkema, Pinkerton & Bloom, 1998;Rabkin & Ferrando, 1997). The authors also suggest that understanding how people cope with HIV/AIDS is important. Psychologists could intervene not only when persons are newly diagnosed but also when new drug therapies bring them back to relatively good health after periods of serious illness. Closely tied to the issue of adherence is the need to focus on psychological treatment of continued risk behavior. If persons with HIV/AIDS who are on drug therapies have unprotected sex with a drug naive HIV/AIDS infected person; they risk infecting their partner with their own strain of virus, which carries its own mutations and possible resistance to drugs. Lastly, Kelly et al.(1998) state the need for behavioral research in prevention and care policy areas.

Clearly, psychology's contribution to prevent new infections of HIV as well as bettering the prognosis of persons diagnosed with HIV or AIDS could be substantial. Psychology continues to search for venues where it can show that its treatments are efficacious and the HIV/AIDS pandemic offers the opportunity of such an examination on a global scale. The face of HIV/AIDS has changed dramatically. Along with these changes is the opportunity for psychology to contribute to the well being of those diagnosed with HIV/AIDS (Kelly, et al, 1998) and the pandemic overall.

Research has shown consistently that training on HIV/AIDS is an issue that needs to be addressed (Anderson, et al., 1998; Campos, et al., 1987; Diaz & Kelly, 1991; Hunt, 1996; Pingitore & Morrison, 1991). The present study will address the following questions: (1) how the degree of specific HIV/AIDS training differs from that reported in the Campos et al. study, (1989) (2) what is the status of courses offered which could easily include HIV/AIDS training (e.g. human sexuality, community psychology) compared to 1987, (3) how much has research increased on HIV/AIDS by students and/or faculty since 1987, (4) what relationships exist between faculty HIV/AIDS research activities and student research involvement in this area, (5) what are the overall importance ratings that psychology place on HIV/AIDS training and (6) what are the barriers to training described by chairs and directors of clinical training.

METHOD

Participants

"Graduate Study in Psychology" (APA, 1994) was used to obtain a listing of 355 psychology departments that offer clinical, counseling or experimental doctoral degrees in areas where HIV/AIDS research could be expected (e.g. Behavioral Medicine/Health Psychology, Biopsychology, Social Psychology, General Psychology, Developmental Psychology, Personality, Cognitive Psychology).

Materials

The original instrument used by Campos et al. (1989) was obtained and adapted for use with research based psychology programs as well as clinical and counseling programs. The 18 item survey included questions about the program (degrees offered, location, number of students and faculty), faculty's research and clinical activities and student training regarding HIV/AIDS. The programs surveyed were also asked to describe how other HIV/AIDS relevant courses (human sexuality, health psychology, substance abuse, primary prevention, minority mental health and community psychology) were presented. In the 1987 survey content areas related to AIDS were assessed as either "not offered", "part of another course", "entire course" or "colloquia". Only 3% of respondents reported content areas being provided as colloquia. Therefore, the present survey assessed whether or not content areas were "required" for graduation in lieu of "colloquia". Four additional items were added to the Campos et al. (1989) instrument to acquire feedback concerning obstacles believed to hinder research in the area of HIV/AIDS, and importance ratings of HIV/AIDS training for future clinicians. To keep the survey to a one-page format and increase response rate, individual questions addressing training on various high risk groups (bisexuals and gays, intravenous drug users and heterosexuals) were collapsed into one category or "at risk groups."

Procedure

A total of 355 cover letters and surveys were mailed to psychology programs. Surveys for research based psychology programs were mailed to the department chair, for clinical programs to directors of clinical training and to the program directors for counseling programs. Two mailings were made. The first half of the surveys were mailed in January (178) and the second in June (177). Programs included in the first mailing, which did not respond, by March , 1995 were sent a second survey and cover letter. A final mailing to non-responding programs was completed in May, 1995. Programs which did not respond to the second mailing were sent a second and last survey and cover letter in August.

RESULTS AND DISCUSSION

Of the 355 surveys mailed, 202 or 57% were returned. Of the responding programs, 57 (28%) were from clinical Ph.D. and 16 (8%) from Psy.D. Psychology programs. Six (3%) surveys were returned from programs that offered both the clinical Ph.D. and Psy.D. degrees. Twenty-seven (14%) from counseling psychology Ph.D. programs, 75 (37%) from experimental Ph.D. programs and 21 (10%) from programs which offered both clinical and experimental psychology Ph.D. degrees. Forty-seven percent of the responses came from programs which were located in large cities, 13% came from programs near but not in large cities, 23% from programs located in medium sized cities and 17% from programs located in small towns.

HIV/AIDS Training in Graduate Program The results of this second assessment of HIV/AIDS training show that progress has been made in training graduate students in psychology regarding HIV and AIDS. In 1987, 75% of the responding clinical programs offered no training on

HIV/AIDS. In 1995, 31% of clinical programs still offer no training on HIV/AIDS, however, this is a marked improvement. Forty-four percent of experimental doctoral programs offer no HIV/AIDS information. Table One shows HIV/AIDS specific training areas in responding graduate psychology programs. Sixty-four percent of the responding programs reported providing HIV/AIDS related information as "part of other courses". And the remaining programs offered HIV/AIDS related information as colloquia (13%) or entire courses (2%). Forty percent of the programs surveyed in 1995 did not offer any training on clinical services or with the HIV/AIDS population. The remaining programs offered training on the clinical aspects of HIV/AIDS either by colloquium (7%) or entire courses (1%). Overall, graduate programs in psychology are offering training in areas that could include HIV/AIDS information and sensitize researchers and clinicians to populations affected by this disease.

TABLE ONE Graduate Student Training in Specific AIDS-Related Content Areas

	Area topic offering			
	NO	P of C	EC	C
Area	%	%	%	%
General HIV Information	31	64	2	13
Services to HIV/AIDS Patients	40	61	1	9
Services to Homosexuals	20	71	6	3
Services to "at risk" populations	35	58	1	6

Note: These data include clinical Ph.D., Psy.D. , joint Ph.D. and Psy.D., counseling Ph.D., experimental Ph.D. and joint experimental and clinical programs. NO = not offered, P of C = part of course, EC = entire course, C= colloquium.

Table Two shows that twenty-percent of programs did not train their students on counseling with the homosexual population. The majority of programs (61%) provide clinical service information with homosexuals as part of other courses. Thirty-five percent of programs did not train psychologists in counseling procedures with at risk populations, such as minorities, drug users and adolescents. The majority of programs (58%) offered information on "at risk" populations as part of other course.

Graduate Training in Other HIV/AIDS Related Areas - Likewise, training in graduate programs that could contain HIV/AIDS related information has increased. Table Two shows how other content related to HIV/AIDS, or courses that could provide HIV/AIDS information, are presented by graduate programs in psychology. Some progress is noted in the fewer number of courses "not offered" which could provide HIV/AIDS information. Campos et al. (1989) reported that only one third of the responding graduate programs in psychology offered courses in human sexuality and 51% offered no human sexuality training. Approximately 28% of the surveyed clinical graduate programs in psychology still do not offer courses on human sexuality. Drug abuse and

TABLE 2 Graduate Student Training in Content Areas Relevant to HIV/AIDS

	Area topic offering			
	NO	P of C	EC	R
Area	%	%	%	%
Human Sexuality	28	35	1	7
Health Psychology	18	22	48	12
Drug Abuse	12	47	35	6
Primary Prevention	19	65	9	7
Minority Mental Health	7	41	32	24
Community Psychology	20	31	38	11

Note. These data include clinical Ph.D., Psy.D. , joint Ph.D. and Psy.D., counseling Ph.D., experimental Ph.D. and joint experimental and clinical programs. NO= not offered, P of C = part of course, EC = entire course, R = required.

primary prevention were also reported by Campos et. al. (1989) to be neglected areas of training, citing 42% of the respondents not offering any courses in these areas. Primary prevention and drug abuse courses were absent in only 12% and 19% respectively of the responding clinical programs. Approximately 47% to 65% of the responding programs report dealing with these subjects as part of other courses. In the areas of minority mental health and community psychology 7% and 20% respectively of programs do not offer any courses in these areas. In the area of health psychology 18% of the clinical programs which responded did not offer any courses in this area. However, approximately 48% offer an entire course dedicated to health psychology. There were some differences in courses provided by clinical and experimental programs. However, this is to be expected given the focus of each type of program and beyond the scope of the present survey.

Graduate Student and Faculty Activities Regarding HIV/AIDS What is also encouraging is that more students are involved in research regarding HIV/AIDS as is shown by the increase in the number of theses and dissertations which are HIV/AIDS related. A total of 17,556 students were enrolled in the responding clinical and experimental programs.

Over the previous two years, 2,927 master's theses and 3,667 doctoral dissertations were completed. Of the 2,927 master's theses completed, 98 (3%) and 158 (4%) of dissertations dealt with HIV/AIDS. This represents a substantial increase to the .4% of HIV/AIDS theses and 1% of doctoral dissertations reported by Campos et al. (1989). It should be noted that the present study examined both research and clinical psychology programs whereas the original study only surveyed clinical and counseling programs. Thus these figures are not directly comparable. However, they do provide a more complete view of psychology doctoral training regarding HIV/AIDS overall.

Examining only clinical programs in 1995 found that of the 2,461 dissertations completed in clinical programs (including clinical Ph.D., Psy.D., counseling Ph.D. and joint experimental and clinical Ph.D. programs), 116 (5%) dealt with HIV/AIDS versus the 1% cited in 1989. Campos et al. (1989) reported that 27 dissertations regarding HIV/AIDS were in progress in 1987. The present survey found 80 HIV/AIDS related dissertations were reported as in progress in clinical programs. The dissertations being produced by graduate students in clinical psychology programs has also increased significantly.

A similar pattern of improvement was found in relation to master's theses. The present survey found that 1,469 theses were reported by clinical programs with 63 (4%) being HIV/AIDS related. Twelve theses were reported to be in progress that addressed HIV/AIDS in the 1987 survey. As of 1995, fifty-one theses were reported to be in progress that focused on HIV/AIDS. In the 1989 survey 932 master's theses were completed. Of those .4 (0.4%) dealt with HIV or AIDS.

The experimental programs surveyed reported similar numbers of HIV/AIDS dissertations (4%) to the clinical program's (5%). However, the number of HIV/AIDS related theses reported by experimental programs was (2%) versus the (4%) reported by clinical doctoral programs. Fifty HIV/AIDS related dissertations and 36 theses were reported as in progress by the experimental programs. There was no statistically significant difference between clinical and experimental programs and mean number of HIV/AIDS related theses or dissertations.

What has not changed substantially is the number of faculty involved in HIV related activities such as research or providing psychological services to persons with HIV/AIDS. A total of 3,438 faculty members were involved in both the clinical and experimental programs. Of these, 112 (3%) of faculty members were conducting research related to HIV/AIDS, 40 (1%) have extramurally funded grants for HIV/AIDS research, 133 (4%) provide clinical or counseling services to persons with HIV/AIDS and 123 (4%) report being involved in other HIV/AIDS related activities. In the previous study 2% of faculty were conducting

HIV/AIDS research, .6% had grants for HIV/AIDS research, 5% were providing clinical or counseling services to persons with HIV/AIDS and 4% were involved in other HIV related activities. As this data shows, very little change has taken place in faculty involvement in HIV/AIDS related activities including clinical services and research with this population. Independent sample *t*-tests were conducted to compare mean differences of number of faculty members involved in HIV/AIDS activities between clinical and experimental programs. Experimental programs had significantly more faculty members involved in HIV/AIDS research (t (188) = -1.25, $p < .05$). Likewise, experimental also had more faculty members working on HIV/AIDS related grants (t (175) = -2.19, $p < .05$).

Correlations between Faculty and Student Activities In general, the best predictors of future theses/dissertations in both clinical and experimental doctoral programs regarding HIV/AIDS are faculty activities in this area. The number of faculty involved with research in the area of HIV/AIDS was correlated with the number of theses in progress in this same area ($r = .61$, $p < .01$). The number of faculty with extramurally funded grants for HIV/AIDS related research was also correlated with the number of theses in progress in this area ($r = .53$, $p < .01$). Similarly, the number of faculty involved in "other" HIV/AIDS related activities was correlated with the number of theses in progress (r =. 43, $p < .01$). The same pattern of correlations were found with dissertations in progress. The number of faculty involved in HIV/AIDS research was the best predictor of the number of dissertations in progress regarding HIV/AIDS ($r = .42$, $p < .01$). Faculty involved with "other" HIV/AIDS activities was also significantly related to the number of future dissertations ($r = .30$, $p < .01$). The number of faculty with HIV/AIDS related grants also correlated with the number of dissertations in progress ($r = .28$, $p < .01$). Lastly, the number of faculty involved in providing clinical services to persons with HIV or AIDS was related to the number of planned dissertations ($r = .20$, $p < .01$).

Importance Ratings of HIV/AIDS Training One hundred ninety-seven of the programs surveyed responded to the question "how important do you feel training regarding HIV/AIDS is for future clinicians?" Thirty-eight percent of the respondents view training regarding HIV/AIDS as very important. Another 23% view HIV/AIDS training as somewhat important. Thirty-seven percent indicated that HIV/AIDS training was important and 2% of respondents felt HIV/AIDS training was not important.

HIV/AIDS Training Issues Of the 202 responses, only 108 (53%) answered the open ended questions added to the original instrument. In response to the question "In what area do you feel psychology can make the greatest contribution to curb the repercussions of the HIV/AIDS?", 17 responses varied from the need for understanding the psychological sequelae of HIV/AIDS to comments that the HIV/AIDS pandemic is not as severe as the Influenza A pandemic.

Ninety-two programs (45%) responded to the question "What do you feel is the greatest obstacle to overcome regarding research and training regarding HIV/AIDS?" A large number of responses 24 (26%) cited funding as a key obstacle. Another 20 (22%) responses cited fear, ignorance and the stigma surrounding HIV/AIDS as a key obstacle to research and training. Interestingly, approximately 9 (10%) respondents focused on the lack of training in graduate programs regarding HIV/AIDS and sexuality as well as the need for recruitment of new faculty and students with an interest in this area. Eight other respondents (9%) cited the need for a better understanding of psychology's contribution to health issues and the need for sound research as an obstacle to overcome regarding training and research with HIV/AIDS. Similarly, eight respondents cited the lack of accessibility of the HIV/AIDS population for research as another obstacle. Seven (8%) responding programs identified that a key obstacle to training and research regarding HIV/AIDS was finding space to add to the already full curricula required by APA. Another seven programs cited politics as an obstacle to HIV/AIDS training and research. Five programs (5%) reported that homophobia and discomfort with sexual and social class issues was a major obstacle to training and research.

Faculty could be more reluctant to train and supervise research in areas where they themselves have not received training. A continuing education course is available for clinicians on the psychological aspects of HIV/AIDS as well as providing general information about this disease. However, research/experimental faculty are not required to take continuing education courses. As would be expected, there was a significant relationship between faculty activities and the number of theses and dissertations being produced. Faculty involvement in HIV/AIDS related activities are also key to securing grants for research in this area, as only persons holding doctorate degrees are usually needed as principal investigators. Clearly, the lack of involvement of faculty members in HIV/AIDS related activities needs to be addressed. The need for faculty involvement was also recognized by the respondents of the present survey. Ten percent of the respondents reported that a lack of HIV/AIDS training and the need for faculty recruitment as an obstacle in this area.

One limitation of the current survey is the inability to infer how the non-responding programs would change the results. Also, the sampling frame has changed due to the addition of new APA accredited programs.

Thus broad generalizations of the present findings are not warranted. HIV/AIDS information continues to change rapidly. New research is helpful in developing new medical treatments for HIV and AIDS, and adding to our knowledge on retroviruses and immunology. Psychologists must be able to keep pace by integrating and disseminating this information. Research has shown that HIV/AIDS training is related to perceived competence in clinical doctoral level students (Kindermann, Morales, & Matteo, 1993) as well as increasing knowledge regarding the psychological aspects of HIV in social workers (Despotes, Cook, Razzano, Grey, Wolf, Linsk, & Mitchell, 2002). APA (1999) has created a fine instructional guide that provides strategies for including HIV/AIDS in various courses, suggestions for the creation of new courses and seminars as well as discussion topics for various courses. The use of existing resources as well as the active recruitment of faculty involved in HIV/AIDS research is necessary for psychology to contribute fully to the HIV/AIDS pandemic.

REFERENCES

American Psychological Association. (1994). *Graduate study in psychology.* Washington, DC: Author.

Anderson, J., Campos, P. & Hamid, G. (1998). *A survey of graduate faculty teaching HIV/AIDS courses.* Washington, DC: American Psychological Association.

Campos, P.E., Brasfield, T.L., & Kelly, J.A. (1989). Psychology training related to AIDS: Survey of doctoral graduate programs and predoctoral internship programs. *Professional Psychology: Research and Practice, 20,* 214-220.

Centers for Disease Control (1998). Combating complacency in HIV prevention; Retrieved from www.the body.com.

Chesney, M. (1993). Health psychology in the 21st century: Acquired immunodeficiency syndrome as a harbinger of things to come. *Health Psychology, 12,* 259-268.

Cohen, S. & Herbert, T.B. (1996). Health psychology: Psychological Factors and physical disease from the perspective of human psychoneuroimmunology. *Annual Review of Psychology, 47,* 113-142.

Cole, J.D., Watt, N.F., West, S.G., Hawkins, J.D., Asarnow, J.r., Markman, H.J., Ramey, S.L., Shure, M.B., & Long, B. (1993). The science of prevention: A conceptual framework and some directions for a national research program. *American Psychologist, 48,* 1013-1022.

Despotes, J., Cook, J., Razzano, L., Grey, D., Wolf, M., Linsk, N.L. & Mitchell, C.G.(2002). Evaluating HIV mental health training: Changes in practice and knowledge for social workers and case managers. *Health & Social Work, 27,* 67-70.

Diaz, Y.E. & Kelly, J.A. (1991). AIDS-related training in U.S. schools of social work. *Social Work, 36,* 38-42.

Ferrando, S. (1998). Behavioral Research on AIDS-Protease Inhibitors and the new millennium: Comment on Kelly, Otto-Salaj, Sikkema, Pinkerton & Bloom. *Health Psychology, 17*, 307-309.

Hunt, B. (1996). HIV/AIDS training in CACREP-Approved counselor education programs. *Journal of Counseling & Development, 74*,295-299.

Kelly, J.A. & Murphy, D.A. (1992). Psychological interventions with AIDS and HIV: Prevention and treatment. Special Issue: Behavioral Medicine: An update for the 1990's. *Journal of Consulting and Clinical Psychology, 60*, 576-585.

Kelly, J.A., Murphy, D.A., Sikkema, K.J., & Kalichman, S.C. (1993). Psychological interventions to prevent HIV infection are urgently needed: New priorities for behavioral research in the second decade of AIDS. *American Psychologist, 48*, 1023-1034.

Kelly, J.A., Otto-Salaj, K.L., Sikkema, K.J., Pinkerton, S.D. & Bloom, F.R. (1998). implications of HIV treatment advances for Behavioral Research on AIDS: Protease inhibitors and new challenges in HIV secondary prevention. *Health Psychology, 17*, 310-319.

Kindermann, S.S., Morales, E. & Matteo, T.M.(1993). HIV training and perceived competence among doctoral students in psychology. *Professional Psychology: Research & Practice, 24*, 224-227.

McDaniel, S.J., Purcell, D.W., Farber, E.W., Bondurant, A., Donovan, J.E., Chang, B. M. & Campos, P.E. (1998). *Academic Psychiatry, 22*, 107-116.

Morton, O. (1992). Achievements in research. In J. Mann, D.J.M. Taranatola & T.W. Netter (Eds.), *A Global Report: AIDS in the world* (pp.229-277). Cambridge, MA; Harvard.

Michaelson, R. (1993, October). Research is underused as weapon in AIDS fight. *APA Monitor*, p.4.

Pingitore, D.P. & Morrison, A. (1993). AIDS-related activities offered by doctoral psychology programs: Results from a national survey. *Professional Psychology: Research and Practice, 24*, 110-114.

Rabkin, J. G. & Ferrando, S. (1997). A "Second Life" agenda: Psychiatric research issues raised by protease inhibitor treatments for people with the human inmunodeficiency virus or the acquired immunodeficiency syndrome. *Archives of General Psychiatry, 54*, 1049-1053.

Wong, F.Y., Duffy, K.G., Faulring, C., Eggleston, B., Harper, G.W. (2001). A content analysis of HIV/AIDS information in psychology textbooks: Implications for education, training and practice. *AIDS Education & Prevention, 13*, 561-570.

Note: The author wishes to thank Loma Linda U. and Mississippi State U. for their financial support and Dr. Hector Betancourt, Dr. Kelly Morton and Dr. Jeffrey Kelly for their input and guidance on earlier versions of this article.

Attachment Styles, View of Self and Negative Affect

Amy Van Buren
Sacred Heart University

Eileen L. Cooley
Agnes Scott College

We investigated the relationship between attachment styles and negative affect using Bartholomew and Horowitz's (1991) model of attachment. Attachment styles with a negative self view (i.e., preoccupied and fearful) were expected to be associated with more distress, especially the fearful style which involves negative views of both self and others. Measures of attachment, depression, depression proneness, and social anxiety were administered to 293 undergraduates. As predicted, participants with "negative self" attachment styles reported more symptoms of depression, proneness to depression, and social anxiety, but, contrary to prediction, those with a fearful style did not report more symptoms of depression and anxiety than those with a preoccupied style. Results suggest that the negative view of self significantly predicts depression and anxiety. Preoccupied and fearful attachment styles may best be described as predicting general negative affectivity. Implications for counseling are discussed.

Attachment theory and concepts were originally developed to address the relationship between infants and their primary caregivers (Bretherton, 1991). However, in the1980s attachment research shifted from a primary focus on the adult-infant relationship to the application of attachment concepts to adult-adult relationships. Early attachment experiences with caregivers are believed to form prototypes or internal working models that set the stage for patterns and expectations in later adult relationships (Berman & Sperling, 1994; Rothbard & Shaver, 1994). These internal working models incorporate a view of the self as loveable or not and a perspective of others regarding whether others are likely to meet an individual's needs or provide rejection. These internal working models function as templates or attachment styles, influencing people's behaviors as they interact with and develop adult relationships, thus providing continuity between child and adult relationship patterns (Rothbard & Shaver, 1994).

Author info: Correspondence should be sent to: Dr. Amy Van Buren, Psych. Dept., Sacred Heart University, 5151 Park Ave., Fairfield, CT, 06432.

North American Journal of Psychology, 2002, Vol. 4, No. 3, 417-430.
© NAJP

Hazan and Shaver (1987) proposed a three-category model of adult attachment by applying Ainsworth's three–part model of infant attachment to describe romantic relationships in adulthood. They demonstrated that secure, anxious-ambivalent, and avoidant styles of attachment were each present in adults who differed in their histories of attachment and their cognitive, internal working models. Securely attached adults were found to perceive love positively but realistically and to view themselves as easy to get to know and like. Individuals with an avoidant style of attachment were found to fear interpersonal closeness. People with an anxious-ambivalent type of attachment were more likely to describe love with concepts of jealousy and obsessiveness and often felt less confident and misunderstood in interpersonal relationships.

Subsequently, Bartholomew and Horowitz (1991) replaced the three-category model with a four-category model of attachment. In their conceptualization of adult attachment, the view of self as positive or negative was crossed with the view of other people as positive or negative. This created four subtypes of attachment and the avoidant classification of attachment was now subdivided into two types. In the four-category model, there are secure, preoccupied, fearful and dismissive types of attachment. The securely attached people see themselves favorably and believe that other people will be responsive to them. For the fearful type, both views of self and others are negative; the person doesn't feel loveable and believes others will be rejecting and untrustworthy. For the preoccupied type (the anxious-ambivalent category in Hazan and Shaver's scheme) the person holds a positive view of others but a negative self-perspective and thus may be more "preoccupied" with relationships, i.e., often obtaining a sense of self by being valued by other people. For the dismissive type, the person has a positive view of the self but a negative view of others and thus may not seek or value relationships.

Using different models of attachment, security of attachment has been studied in relation to general psychopathology and negative mood states. Thoughts and behaviors of people with a secure style of attachment have been compared to those of people with insecure styles. For example, Mikulincer and Florian (1998) examined the coping ability of adults under stressful situations and found an insecure style of attachment was associated with more negative responses to stress, while securely attached adults managed stress more effectively and positively. In a study of college students, Kemp and Neimeyer (1999) also noted that attachment styles were associated with reactions to stress. Specifically, preoccupied attachment was correlated with more reports of negative symptoms and a feeling of distress. Administering the MMPI-2 to a low-income sample of first-time mothers, Pianta, Egeland and Adam (1996), using the Adult Attachment Interview, similarly found that attachment style predicted psychiatric symptoms; women with a preoccupied style reported the most symptoms of distress.

Other investigations of attachment and psychological distress have been applied to specific mood states, especially depression. Many studies have focused on student samples using both three-category and four-category models of attachment. Using the three-category model Roberts, Gotlib, and Kassel (1996) found that higher scores on a measure of depressive symptomatology were associated with both the anxious/ambivalent and avoidant styles, but in a regression equation only the anxious ambivalent style made a specific contribution to symptoms of depression. Using the same model of attachment, Priel and Shamai (1995) found that anxious/ambivalent students had more symptoms of depression than avoidant students. Using Bartholomew and Horowitz's four-category model, Carnelley, Pietromonaco, and Jaffe (1994) found that depression in college women was associated with both preoccupied and fearful styles of attachment.

Insecure styles of attachment also have been associated with depression in adolescent and adult samples. Interviewing psychiatrically-hospitalized adolescents with the Adult Attachment Interview, Rosenstein and Horowitz (1996) found that preoccupied and dismissive styles of attachment were correlated with specific diagnoses. A diagnosis of affective disorder was associated with a preoccupied style of attachment, while a dismissing style was associated with conduct disorder. For married women with a history of clinical depression, responses to Bartholomew and Horowitz's model revealed that a fearful, and not preoccupied, style was associated with depression (Carnelley et al., 1994). Similarly, in a study of adult couples, Whiffen, Kallos-Lilly, and MacDonald (2001) reported that depressed women were more likely to have a fearful style of attachment than a comparison sample.

These findings on the relationship between depression and attachment are consistent with Beck's (Beck, Rush, Shaw, & Emery, 1979) cognitive triad model of depression. He proposed that people get depressed because of their negative view of the self, experience and the future. Based on Beck's conceptualization, insecure attachment styles incorporating a negative view of the self (i.e., preoccupied or fearful) would be associated with more depressive symptoms than attachment styles with a positive view of the self (i.e., secure or dismissive) (Carnelley et al., 1994). However, there appear to be inconsistent findings regarding the particular style of attachment that is most often linked with symptoms of depression. Some research suggests the anxious/ambivalent or the preoccupied style) is most vulnerable (e.g., Priel & Shamai, 1995; Rosenstein & Horowitz, 1996), while other studies propose that the fearful style best predicts depressive symptoms (e.g., Carnelley et al., 1994; Whiffen et al., 2001). Carnelley, et al., (1994) found that both preoccupied and fearful styles predicted depressive symptoms with a college student sample. This situation is further complicated since these studies used different models of attachment and varied in their use of adolescent, adult, or college student samples. In Hazan and Shaver's (1987) three-category model, the avoidant type includes people with both positive and negative views of themselves. Studies based on this model, therefore, don't directly test the fearful style of attachment and their results of specific attachment styles are then more difficult to interpret. As current research emphasizes the four-category model of attachment, more research using this model is needed to clarify the relationship between particular types of attachment styles and depression.

Furthermore, it remains unclear whether the association between depression and attachment represents a finding specific to depression or a more general finding for negative mood states or distress. For example, few studies have examined the relationship between attachment style and anxiety. Priel and Shamai (1995) used the State form of the *State-Trait Anxiety Inventory* in their study of college students (three-category model of attachment) and found that anxious-ambivalent students indicated more anxiety than avoidant students who, in turn, reported more anxiety than secure students. Leondari and Kiosseoglou (2000) studied university students from Greece using the *Beck Anxiety Inventory* and the *Inventory of Parent and Peer Attachment*. They found that security of attachment predicted anxiety and loneliness but they did not measure specific styles of attachment. It seems likely that participants with a negative view of themselves would also be more likely to report higher levels of anxiety, as they might worry about their performance, especially in their interpersonal interactions. However, attachment research incorporating specific types of attachment styles is needed to examine the relationship between anxiety and attachment.

The present study sought to refine our understanding of the relationship between styles of attachment, and depression and anxiety. Previous research suggests that a negative view of the self is likely to be associated with symptoms of depression. We expected to replicate this. However, prior research has produced mixed findings regarding which attachment style with a negative view of the self is most likely to be associated with depression (i.e., preoccupied or fearful). Based on Beck's theory of depression and Bartholomew and Horowitz's (1991) model of attachment, our first hypothesis was that that the fearful style, which includes a negative view of the self and others, would be most likely to be associated with symptoms of depression. The fearful style of attachment represents a more extreme form of negativity that goes beyond the self, thus expanding the person's negative perspective and experience.

In addition, we examined proneness to depression to determine if the relationship between attachment styles and symptoms of depression could be extended to include a trait perspective of depression. Most prior research has focused on immediate depressive symptoms rather than long-term patterns of depression. Our second hypothesis was that people reporting negative styles of attachment would be more likely to report being vulnerable to depressive symptoms throughout their lives.

Our third hypothesis was that attachment styles with a negative view of the self, particularly the fearful style, would be associated with other negative mood states, specifically interpersonal anxiety. Since people with the fearful style of attachment were found to be socially inhibited and lacking in assertiveness skills (Bartholomew & Horowitz, 1991), they also were expected to report higher levels of interpersonal anxiety. An attachment style incorporating a negative view of the self and others may be associated with negative mood states in general, rather than specifically linked to depression.

In summary, it was hypothesized that participants with a negative view of themselves (i.e., preoccupied and fearful) would report more symptoms of depression and interpersonal anxiety as well as a greater proneness to depression than participants with a positive view of themselves (i.e., secure and dismissive). Furthermore, it was predicted that participants with a fearful style of attachment would indicate the highest levels of depression, interpersonal anxiety, and proneness to depression.

METHOD

Participants

Participants were 293 undergraduates attending either a small liberal arts college for women in the southeast (*n* = 149) or a private coeducational liberal arts college in the northeastern United States (*n* = 144). Eighty-eight percent were women and 12 % were men. Regarding ethnic identity, 78 % were Caucasian, 12% African-American, 2% Hispanic, 3 % Asian-American, and 5% other. Ninety-five percent of the participants were single, 2% divorced or separated, and 3% married. Of the participants who were single, 60% were not dating or dating casually while 40% reported they were in a serious relationship. Data were collected over a two-year period as part of two separate studies on social relationships and attachment. Therefore the number of participants varies for the various measures in this study.

Measures

The Relationship Styles Questionnaire (RSQ; Bartholomew & Horowitz, 1991). The RSQ is based on a 4-group model of attachment proposed by Bartholomew and Horowitz. It consists of four short paragraphs, each of which describes one of the four attachment styles. Participants are asked to indicate which of the four descriptions is most characteristic of them. Empirical data support the utility of the four-group model (Bartholomew & Horowitz, 1991).

Beck Depression Inventory II (BDI-II; Beck, Steer, & Brown, 1996). The BDI-II is a 21 item self report scale measuring current depression. Each item is rated on a 0-3 scale, with higher scores corresponding to more symptoms of depression. In this study, item #9 assessing suicidal ideation was omitted. This was done after a faculty reviewer at one of the institutions expressed concern that given the anonymity of the study, there would be no way to follow up with a participant who answered the item in the affirmative. Regarding reliability, the coefficient alpha for 120 college students was .93 and the test-retest correlation for 26 outpatients was .93 after a one-week interval (Beck et al., 1996). The BDI-II is positively related to the *Beck Hopelessness Scale* (r = .68) and to the *Hamilton Psychiatric Rating Scale for Depression* (r = .71), indicating adequate validity (Beck, et al., 1996).

The Depression Proneness Rating Scale (DPRS; Zemore, 1983). The DPRS is a 13-item scale designed to assess a person's tendency to become depressed. Participants read each statement and rate each depressive symptom for frequency of occurrence over the past two years. Each item is rated on a 9-point scale with higher numbers indicating greater proneness to depression. Coefficient alpha calculated from the original subject pool of 360 students, was .83. Two-week, four-week and six-week test-retest correlations were .90, .82 and .72, respectively (Zemore & Dell, 1983). As reported in Zemore (1983), students' scores on the DPRS correlated significantly with depression-proneness ratings by their parents and peers.

Interaction Anxiousness Scale (IAS; Leary, 1983). The IAS is a 15 item self report measure of the tendency to experience subjective social anxiety independent of accompanying behaviors. Items consist of 15 statements that are responded to on a 5-point scale ranging from "not at all characteristic of me" to "extremely characteristic of me." In terms of internal consistency, all items correlate at least .45 with the sum of all other items and Cronbach's alpha exceeds .87. Eight week test-retest reliability is .80 (Leary, 1991). IAS scores correlate highly with other measures of social anxiousness and shyness (Jones, Briggs, & Smith, 1986; Leary & Kowalski, 1987). In addition, IAS scores correlate highly with self-reported anxiety in real interactions (Leary, 1983; 1986). For the present study, the title of this measure was changed to the "*Interpersonal Interaction Scale*."

Procedure

The study took place at two locations: a small, private liberal arts college for women in the southeast and a small, liberal arts university in the northeast. Participants at both institutions were recruited through psychology classes and were offered extra credit for taking part in the study by the individual professors teaching those classes.

Participants were tested in small groups and were administered the above measures in a fixed order as part of a larger study. The order was as follows: informed consent; demographic questionnaire measuring age, gender, relationship status and marital status, year in college and major; *Relationship Styles Questionnaire; Beck Depression Inventory-II; Depression Proneness Rating Scale;* and *Interaction Anxiousness Scale.*

RESULTS

Means and standard deviations were computed for each of the target measures. Since the data were collected as part of more than one study on attachment, not all participants completed each of the measures. For the BDI-II there were 238 participants with an overall mean score of 11.19 (*SD* = 8.40), for the DPRS there were 133 respondents with a mean of 54.45 (*SD* = 19.25), while the mean score on the IAS was 41.56 (*SD* = 11.75) for 123 participants.

Attachment responses were coded based on Bartholomew and Horowitz's (1991) model. Forty-four percent of the sample indicated the secure description best fit their behavior in close relationships, 14% the dismissive type, 15% the preoccupied type, and 27% the fearful type. Participants reporting a secure or dismissive style of attachment were identified as "positive self" (58%) and those indicating a preoccupied or fearful style were labeled as "negative self" (42%).

Using the positive v. negative "self" groups, an independent t-test compared BDI-II scores between groups to test the hypothesis that people with a negative view of themselves would report more symptoms of depression than people with a positive self view. As predicted the "negative self" styles (i.e., fearful or preoccupied)), were associated with significantly higher BDI-II scores (M = 14.04, SD = 9.17) than those with a "positive self" view (i.e. secure or dismissive) (M = 8.98, SD = 7.02), $t_{(236)} = -4.82$, $p < .001$, (Cohen's effect size $d = .63$, (Cohen, 1988).

To examine whether the fearful style of attachment was most likely to be associated with symptoms of depression, a one-way ANOVA was conducted using attachment style as the between subject's variable and BDI-II scores as the dependent variable. The mean scores for the secure, dismissive, preoccupied and fearful styles of attachment were: 8.32 (6.79), 10.91 (7.43), 15.05 (9.37), and 13.48 (9.09). The ANOVA was statistically significant, $F_{(3, 237)} = 8.99$, $p < .001$ (Cohen's effect size $\underline{f} = .34$ (Cohen, 1988) and Scheffe's post-hoc test showed that both the preoccupied (mean difference = -6.73, $p < .01$) and fearful (mean difference = -5.16, $p < .01$) attachment groups reported significantly more depressive symptoms on the BDI-II than the secure group. Contrary to prediction, these two groups were not significantly different from each other.

Depression-proneness, a trait perspective on depression, was analyzed in the same way as the BDI-II. The group with a "negative self" view reported more proneness to depression ($M = 63.25$, $SD = 16.81$) than the group with a "positive" view of the self, ($M = 48.24$, $SD = 18.52$) t (131) = - 4.78, $p < .001$, (effect size $d = .84$). A one way ANOVA with attachment style as the between subject's factor and DPRS as the dependent variable was statistically significant, $F_{(3,132)} = 8.51$, $p < .001$ (effect size $f = .45$). Scheffe's post-hoc test showed that both the preoccupied (mean difference = -19.85, $p < .01$) and fearful (mean difference = -12.74, $p < .02$) attachment groups reported more proneness to depression than the secure group, but were not significantly different from each other. In addition the dismissive group had statistically lower DPRS scores than the preoccupied group of participants (mean difference = -16.63, $p < .05$). The means for the specific attachment types were as follows: secure 47.54 ($SD = 17.83$), dismissive 50.76 ($SD = 21.22$), preoccupied 67.39 ($SD = 21.31$), and fearful 60.28 ($SD = 12.16$).

For interpersonal anxiety, the participants with a negative view of the self ($M = 47.24$, $SD = 12.63$) produced higher scores than participants with a positive view of the self ($M = 38.62$, $SD = 10.15$) on IAS scores, $t_{(121)} = - 4.10$, $p < .001$(effect size $d = .78$). A one way ANOVA with attachment style as the between subject's factor and IAS as the dependent factor was statistically significantly, $F_{(3, 122)} = 6.20$, $p = .001$ (effect size $f = .40$). Scheffe's post-hoc test demonstrated that both the preoccupied (mean difference = -9.62, $p < .02$) and fearful (mean difference = -9.36, $p < .01$) attachment groups reported significantly more anxiety than the secure group and the preoccupied and fearful groups were not significantly different from each other. The means for the specific attachment types were as follows: secure 37.76 ($SD = 9.37$), dismissive 41.82 ($SD = 12.72$), preoccupied 47.39 ($SD = 12.84$), and fearful 47.13 ($SD = 11.75$).

DISCUSSION

Consistent with our hypothesis and Bartholomew and Horowitz's (1991) four-category model, we found that people with a negative view of self (i.e., fearful or preoccupied attachment style) reported more current symptoms of depression than people indicating a positive view of the self (i.e., secure or dismissive style). This finding supports Beck's theory (Beck, 1979) regarding the central role of the self in depression. It is also consistent with previous research (e.g., Carnelley, et al., 1994; Murphy & Bates, 1997) and, given the medium-to-large effect size, suggests a meaningful relationship between a negative self view and current symptoms of depression.

However, the results did not support the prediction that people with the most depressive symptomatology would be those with a fearful style, which involves negative view of *both* self and others. It turns out that participants with a fearful or preoccupied style (which involves a negative view of self but a positive view of others) did not differ from each other in their report of depressive symptoms. It seems that regardless of how one sees others, it is the negative view of oneself that is most consistently linked to symptoms of depression. However, the nature of the link between negative self-view and depression may differ depending on the internal working model of other. For example, preoccupied people who are depressed may show a heightened awareness and reaction to negative social cues, which then affects their view of self and may impair their ability to see themselves objectively (Lopez, 1995). In contrast, those with a fearful style may shut out constructive social feedback or shun potentially helpful relationships, which may then further impair their self-views.

Our findings are consistent with the results of Carnelley, et al. (1994) who used the same four-category model with a college student sample. However, their sample differed from the present sample. The majority of their sample (60%) reported they were currently part of a stable relationship while, in contrast, the majority of our participants indicated they were not in a serious relationship (60%). Taken together, the results of both studies suggest that the opinion one has of oneself plays a greater role in self reported depression than how one feels about others, regardless of whether one is in a serious relationship. For college students, serious relationships may not have continued long enough to provide a corrective experience that may alter the association between depression and attachment style.

Attachment styles with a negative self-concept also were more likely to be associated with a long-term proneness to depression. The pattern of results demonstrated a large effect size and was similar to the pattern reported for current symptoms of depression. Students with an attachment style incorporating a negative self view (i.e., fearful and preoccupied) reported significantly greater proneness to depression than those who had a positive self view (i.e., secure or dismissive) and again, the fearful style of attachment did not differ from the preoccupied style. These findings therefore extend our view of the relationship between depression and attachment styles to incorporate both trait and state experiences of depression. Having a long-term proneness to depression may, in turn, make one more vulnerable to developing acute symptoms. Again, this is true regardless of whether one's view of others leads one to be overly dependent on others (i.e., preoccupied style) or to lack trust in and fear involvement with others (i.e., fearful style).

Finally, we extended the findings on attachment style and depression to include social anxiety. Consistent with our prediction, those whose attachment style consisted of a negative view of self-reported more subjective social anxiety with a medium-to-large effect size. Once again, the fearful style of attachment was not associated with the most interpersonal anxiety, as both fearful and preoccupied styles differed from secure styles, suggesting again that it is a poor self (rather than other) representation that is the critical factor. These findings, however, are important in extending the research on attachment to other negative moods, this time social anxiety.

There are several conclusions that can be drawn from our findings. First, it seems clear that attachment styles with a negative view of self are associated with depression and social anxiety. However, since ours was a college student sample, it is not known whether this would also be true for a clinical sample with diagnosed affective or anxiety disorders. Second, while past research has primarily studied the relationship between attachment styles and depression, we have extended the literature to include the relationship between attachment styles and both trait depression and social anxiety. Our results suggest that having an attachment style with a poor self-view is not specific to current symptoms of depression, but actually involves a cluster of symptoms and personality traits. These include: long-term vulnerability to depression, social anxiety, and anger (Mikulincer, 1998), which together might be better termed "negative affectivity" (Feeney & Noller, 1996; Watson & Clark, 1984; Watson & Pennebaker, 1989).

Developmentally, the relationships among attachment style, self view and negative affectivity are complex. It is likely that biological and temperamental factors may set the initial stage for negative states and traits. During the course of early development, attachment experiences would then further determine whether the person develops an attachment style characterized primarily by a negative self-view. Attachment style and negative affectivity would then mutually influence one another.

However, what is not yet known is the direction of the relationship, i.e., whether attachment style influences the individual's experience of negative affect or whether negative affect influences the kind of attachment style one develops. Further research is needed to investigate this relationship.

Results from the present study have important implications for the counseling process. Our results indicate that the negative view of self is an appropriate target for intervention in people with depression and anxiety, as Beck would suggest. However, when the results are interpreted within an attachment framework, which takes into account view of others as well as of self, it becomes obvious that simply focusing interventions on the self may not be enough. Because fearful and preoccupied people have different self-other schemas, it follows that their interpersonal interactions may be very different. Therefore, an effective counseling approach will differ for each style. For example, preoccupied people, whose symptoms may stem from being other-oriented and overly attuned to subtle and particularly negative social cues, may need help focusing on more objective self-awareness. In contrast, fearful people, whose negative view of others causes them to disengage may need help being more connected to others (Lopez, 1995). Therapy with fearful people may need to involve particular attention to the development of a therapeutic working alliance.

These findings must also be interpreted within a developmental framework. College students are young adults, many of whom have had limited experience in adult-adult romantic relationships. As these students develop and have more extensive relationship experiences, their specific attachment styles may be modified. It would be interesting to follow students during and beyond their college years to catalogue relationship experiences and perhaps changing attachment styles that may accompany their growth and development. In contrast to the present findings, some researchers have found that the fearful style of attachment was most likely to be associated with depression among married women (Carnelley et al., 1994; Whiffen et al., 2001).

There are several factors specific to this study which may limit the conclusions. Since the majority of the participants (88%) were women, the conclusions are more applicable to women than men. Depression in women may be of a more interpersonal nature than depression in men and it may be that attachment models of depression are more applicable to women. If attachment models of depression do apply to men, it is possible that in men, the underlying relationships between self and other are different. If this is so, therapeutic interventions need to differ for the genders. For example, depression in men may be more related to the dismissive (positive self and negative other) style and it follows that counseling approaches will need to be different than for fearful or preoccupied styles. Clearly, more research needs to be done using the four-category model to study attachment and negative mood states in men. Our study is also limited by the fact that the measures are self-report. In future studies, it will be important to gather data measuring long-term depression proneness and interpersonal anxiety from other sources, for example friends and significant others. Finally, it will be important to continue this research by using the four-category model to find whether the self/other distinction applies to other negative states, such as anger and frustration. It will also be important to find out whether the negative view of self continues to be associated with general negative affectivity or only with certain specific negative mood states and traits.

REFERENCES

Bartholomew, K., & Horowitz, L. M. (1991). Attachment styles among young adults: A test of a four-category model. *Journal of Personality and Social Psychology, 61,* 226-244.

Beck A. T. (1979) *Cognitive therapy and depression.* New York: Guilford Press.

Beck, A. T., Steer. R. A., & Brown, G. K. (1996). *Beck Depression Inventory-II Manual.* San Antonio: The Psychological Corporation.

Berman, W. H., & Sperling, M. B. (1994). The structure and function of adult attachment. In M. B. Sperling & W. H. Berman (Eds.), *Attachment in adults: Clinical and developmental perspectives* (pp. 3-28). New York: Guilford Press.

Bretherton, I. (1991). The roots and growing points of attachment theory. In C. M. Parkes, J. Stevenson-Hinde, & P. Marris (Eds.) *Attachment across the life cycle* (pp. 9-32), London: Routledge.

Carnelley, K. B., Pietromonaco, P. R., & Jaffe, K. (1994). Depression, working models of others, and relationship functioning. *Journal of Personality and Social Psychology, 66,* 127-140.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Feeney, J., & Noller, P. (1996). *Adult attachment.* Thousand Oaks, CA: Sage.

Hazan, C., & Shaver, P. (1987). Romantic love conceptualized as an attachment process. *Journal of Personality and Social Psychology, 52,* 511-524.

Jones, W.H., Briggs, S.R., & Smith, T.G. (1986). Shyness: Conceptualization and measurement. *Journal of Personality and Social Psychology, 51,* 629-639.

Kemp, M. A., & Neimeyer, G. J. (1999). Interpersonal attachment: Experiencing, expressing, and coping with stress. *Journal of Counseling Psychology, 46,* 388-394.

Leary, M. (1991). Social anxiety, shyness and related constructs. In J.P. Robinson, P.R. Shaver, & L.S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes: Volume 1 in Measures of social psychological attitudes series.* (pp. 161-194). San Diego, CA: Academic Press.

Leary, M. R. (1986). The impact of interactional impediments on social anxiety and self presentation. *Journal of Experimental Social Psychology, 22,* 122-135.

Leary, M. R. (1983). Social anxiousness: The construct and its measurement. *Journal of Personality Assessment, 47,* 66-75.

Leary, M.R., & Kowalski, R.M (1987). Manual for the Interaction Anxiousness Scale. *Social and Behavioral Sciences Documents, 16,* (Ms. No. 2774), 2.

Leondari, A., & Kiosseoglou, G. (2000). The relationship of parental attachment and psychological separation to the psychological functioning of young adults. *The Journal of Social Psychology, 140,* 451-464.

Lopez, F. G. (1995). Contemporary attachment theory: an introduction with implications for counseling psychology. *The Counseling Psychologist, 23,* 395-415.

Mikulincer, M. (1998). Adult attachment style and individual differences in functional versus dysfunctional experiences of anger. *Journal of Personality and Social Psychology, 74,* 513-524.

Mikulincer, M., & Florian, V. (1998). The relationship between adult attachment styles and emotional and cognitive reactions to stressful events. In J. A. Simpson, & W. S. Rholes (Eds.), *Attachment theory and close relationships* (pp. 143-165). New York: Guilford Press.

Murphy, B., & Bates, G. (1997). Adult attachment style and vulnerability to depression. *Personality and Individual Differences, 22,* 835-844.

Pianta, R. C., Egeland, B., & Adam, E. K. (1996). Adult attachment classification and self-reported psychiatric symptomatology as assessed by the Minnesota Multiphasic Personality Inventory-2. *Journal of Consulting and Clinical Psychology, 64,* 273-281.

Priel, B., & Shamai, D. (1995). Attachment style and perceived social support: Effects on affect regulation. *Personality and Individual Differences, 19,* 235-241.

Roberts, J. E., Gotlib, I. H., & Kassel, J. D. (1996). Adult attachment security and symptoms of depression: The mediating roles of dysfunctional attitudes and low self-esteem. *Journal of Personality and Social Psychology, 70,* 310-320.

Rosenstein, D. S., & Horowitz, H. A. (1996). Adolescent attachment and psychopathology. *Journal of Consulting and Clinical Psychology, 64,* 244-253.

Rothbard, J. C., & Shaver, P. R. (1994). Continuity of attachment across the life span. In M. B. Sperling & W. H. Berman (Eds), *Attachment in adults: Clinical and developmental perspectives* (pp. 31-71). New York: Guilford Press.

Watson, D., & Clark, L. (1984). Negative affectivity: The disposition to experience aversive emotional states. *Psychological Bulletin, 96,* 465-490.

Watson, D., & Pennebaker, J. (1989). Health complaints, stress, and distress: Exploring the central role of negative affectivity. *Psychological Review, 96,* 234-254.

Whiffen, V. E., Kallos-Lilly, A. V., & MacDonald, B. J. (2001). Depression and attachment in couples. *Cognitive Therapy and Research, 25,* 577-590.

Zemore, R. (1983). Development of a self-report measure of depression-proneness. *Psychological Reports, 52,* 211-216.

Zemore, R., & Dell, L.W. (1983). Interpersonal problem-solving skills and depression-proneness. *Personality and Social Psychology Bulletin, 9,* 231-235.

Assessing Attitude Toward Christianity Among Chinese Speaking Adolescents in Hong Kong: the Francis Scale

Leslie J Francis
University of Wales, Bangor, UK

Christopher Alan Lewis
University of Ulster at Magee College, Northern Ireland, UK

Peter Ng
Chinese University of Hong Kong

A sample of 598 Chinese speaking adolescents between the ages of thirteen and eighteen years attending secondary schools within the state-maintained sector in Hong Kong completed the Chinese translation of the Francis Scale of Attitude toward Christianity. The data support the reliability and validity of this instrument and commend it for further use in studies conducted among adolescents in Hong Kong.

The *Francis Scale of Attitude toward Christianity* was developed during the 1970s to provide a reliable and valid measure of attitude toward Christianity appropriate for use among children, adolescents and adults (Francis, 1978). The 24-item Likert scale contains both negative and positive items concerned with an affective response to five components of the Christian faith accessible to and recognised by both children and adults, namely God, Jesus, bible, prayer and church. Each item is assessed on a five-point scale (agree strongly, agree, not certain, disagree, disagree strongly), producing a range of scores from 24 to 120. The reliability and validity of the scale have been supported by studies among school pupils in England (Francis, 1987, 1989), Kenya (Fulljames & Francis, 1987), Nigeria (Francis & McCarron, 1989), Northern Ireland (Francis & Greer, 1990; Greer & Francis, 1991) and Scotland (Gibson, 1989; Gibson & Francis, 1989). Another series of studies have supported the reliability and validity of the scale among adults in Australia and Canada (Francis, Lewis, Philipchalk, Brown, & Lester, 1995), England (Francis & Stubbs, 1987;

Author info: Correspondence should be sent to: Leslie J Francis, Welsh National Centre for Religious Education, University of Wales, Bangor, Bangor, Gwynedd LL57 2PX Wales UK

North American Journal of Psychology, 2002, Vol. 4, No. 3, 431-440.

© NAJP

Francis, 1992), the Republic of Ireland (Maltby, 1994), Northern Ireland (Lewis & Maltby, 1997) and the USA (Lewis & Maltby, 1995).

By the mid-1990s over one hundred independent studies had employed this scale to examine a wide range of correlates of religiosity during childhood, adolescence and adulthood. These studies were summarised and synthesised by Kay and Francis (1996). During the second half of the 1990s the scale has been employed in further studies exploring the correlates of religiosity, including: altruism (Eckert & Lester, 1997); conservatism (Lewis & Maltby, 2000); creationism and evolutionary theory (Francis & Greer, 1999a); denominational differences (Francis & Greer, 1999b); dissociation (Dorahy & Lewis, 2001); extraversion, neuroticism and psychoticism (Youtika, Joseph, & Diduca, 1999); gender orientation (Francis & Wilcox, 1998); happiness (Lewis, Lanigan, Joseph, & de Fockert, 1997); intelligence (Francis, 1998); life satisfaction (Lewis, 1998); obsessionality (Lewis, 1996); operational thinking (Kay, Francis, & Gibson, 1996); personality type (Fearn, Francis, & Wilcox, 2001); prosocial values (Schludermann, Schludermann, & Huynh, 2000) psychological well-being (Francis, Jones, & Wilcox, 1997); purpose in life (French & Joseph, 1999); schizotypal traits (Joseph & Diduca, 2001); and social desirability (Lewis, 2000).

During the second half of the 1990s a series of studies has reported on the successful translation of the *Francis scale of Attitude toward Christianity* into other languages, in order to facilitate the integration of cross-cultural quantitative studies in the psychology of religion through common instrumentation. Examples are provided by editions in Welsh (Evans & Francis, 1996); German (Francis & Kwiran, 1999); Greek (Youtika, Joseph, & Diduca, 1999); Dutch (Francis & Hermans, 2000); Arabic (Munayer, 2000); Swedish (Eek, 2001); and Czech (Francis, Quesnell, & Lewis, 2002).

The present study builds on a recent strand of research concerned with testing the psychometric properties of the Chinese translation of established psychological instruments. For example, Chinese translations have been developed of the *Beck Hopelessness Scale* (Au & Watkins, 1997), the *Bem Sex Role Inventory* (Lau, 1989; Lau & Wong, 1992), the *Coopersmith Self-esteem Inventory* (Chan & Verma, 1995; Chan, 2000), the *Fenigstein Self-Consciousness Scale* (Chan, 1996), the *Harter Perceived Competence Scale for Children* (Stigler, Smith, & Mao, 1985; Chan, 1997), the *Junior Eysenck Personality Questionnaire* (Ng, Cooper, & Chandler, 1998), the *Marsh Self-Description Questionnaire* (Leung & Lau, 1989), the Misiti, Shrigley, & Hanson *Science Attitude Scale* (Boone, 1997), the Moos and Moos *Family Environmental Scale* (Lau & Cheung, 1987), the Moos and Trickett *Classroom Environmental Scale* (Cheung & Lau, 1985), the Nowicki and Strickland *Locus of Control Scale* (Lau, 1990), the *Oxford Happiness Inventory* (Lu & Shih, 1997; Lu, Shih, Lin, & Ju, 1997), the *Rokeach Value*

Survey (Lau, 1988), the *Rosenberg Self-Esteem Scale* (Leung, 1990; Lau & Leung, 1992a, 1992b; Cheung & Kwok, 1996), the Rushton, Chrisjohn, and Fekken *Self-Report Altruism Scale* (Chou, 1996), and the Russell, Peplau, and Cutrona *Revised UCLA Loneliness Scale* (Man & Hamid, 1998). Against this background the aim of the present study was to examine the psychometric properties of the Chinese translation of the *Francis Scale of Attitude toward Christianity* among a sample of Chinese-speaking adolescents.

METHOD

Instrument

The *Francis Scale of Attitude toward Christianity* is a 24-item Likert instrument concerned with affective response toward God, Jesus, bible, prayer and church. The items were translated into Chinese and then back-translated into English in order to test for inaccuracies and ambiguities. Each item is assessed on a five-point scale: *agree strongly*, *agree*, *not certain*, *disagree* and *disagree strongly*.

Church attendance was assessed on a five-point scale: *never*, *once or twice a year*, *sometimes*, *once a week* and *daily*.

Personal prayer was assessed on a five-point scale: *never*, *once or twice a year*, *sometimes*, *once a week* and *daily*.

Sample

The instrument was completed by 598 Chinese-speaking pupils attending secondary schools within the state maintained system in Hong Kong. Of the total respondents 11 were thirteen years of age, 110 were fourteen, 169 were fifteen, 99 were sixteen, 142 were seventeen and 67 were eighteen; 43% were male and 57% were female. Four-fifths of the respondents (79%) described themselves as belonging to no religious group, while the remaining 21% identified a Christian denomination.

Procedure

The *Francis Scale of Attitude toward Christianity* was translated into Chinese by the third author and back translated by a bilingual psychologist.

RESULTS

In the first column, Table One presents the item to rest-of-test correlation coefficients for all 24 items, together with the alpha coefficients. In the second column, Table One presents the loadings on the first factor of the unrotated solution proposed by principal component analysis, together with the percentage of variance explained. Both sets of statistics support the conclusion that the scale is characterised by homogeneity, unidimensionality, and internal consistency reliability within the sample. The alpha coefficient is established as .97, while the proportion of variance accounted for by the first factor is established as 56.9%.

TABLE 1 Item to Rest-of-test Correlations, Alpha Coefficient, and Factor Loadings

Scale item	r	factor loading
I find it boring to listen to the bible*	.48	.50
I know that Jesus helps me	.88	.86
Saying my prayers helps me a lot	.80	.82
The church is very important to me	.65	.68
I think going to church is a waste of my time*	.51	.53
I want to love Jesus	.80	.83
I think church services are boring*	.51	.53
I think people who pray are stupid*	.66	.68
God helps me to lead a better life	.74	.77
I like to learn about God very much	.75	.77
God means a lot to me	.81	.83
I believe that God helps people	.80	.83
Prayer helps me a lot	.82	.84
I know that Jesus is very close to me	.81	.83
I think praying is a good thing	.68	.71
I think the bible is out of date*	.52	.54
I believe that God listens to prayers	.75	.78
Jesus doesn't mean anything to me*	.74	.76
God is very real to me	.76	.79
I think saying prayers does no good*	.72	.74
The idea of God means much to me	.78	.81
I believe that Jesus still helps people	.80	.83
I know that God helps me	.84	.87
I find it hard to believe in God*	.73	.76
alpha coefficient/% of variance	.97	56.9%

Note * these negative items were reverse scored. N = 598.

While the formal statistics of reliability are relatively easy to calculate for an attitude scale (Livingston, 1988), the question of assessing validity is more problematic (Zeller, 1988). Steps toward assessing the construct validity of this scale can be made by assessing the extent to which certain predictions about the theoretical variations in attitude scores are reflected empirically (Orton, 1987). While attitudes alone may not be simple or direct predictors of behaviour (Ajzen, 1988), substantial evidence suggests a fairly close relationship between attitude towards religion and religious behaviour, as demonstrated by repeated studies using the *Francis Scale of Attitude toward Christianity* among children and adolescents (Francis, 1989). In the present study attitude scores correlated with church attendance ($r_{596} = .51$, $p<.001$) and with personal prayer ($r_{596} = .48$, $p<.001$). These statistics support the construct validity of the attitude scale within the sample.

Considerable evidence exists in the empirical study of religious differences to indicate that females record higher scores on indices of religiosity than males (Francis, 1997). The mean scale score for males ($M = 71.7$, $SD = 18.7$) was less than the mean scale score for females ($M = 82.7$, $SD = 16.9$), $t_{(597)} = 7.5$, $p < .001$.

DISCUSSION

The present study has explored the scale properties of the Chinese translation among a sample of 598 young people between the ages of thirteen and eighteen attending secondary schools within the state-maintained sector. The data support the unidimensionality, internal consistency reliability, and construct validity of the Chinese translation of this instrument. The scale can, therefore, be recommended for further use within Hong Kong and should lead to valuable cross-cultural comparisons in an empirical psychology of religion.

REFERENCES

Ajzen, I. (1988). *Attitudes, personality and behaviour*. Milton Keynes: Open University Press.

Au, C.P., & Watkins, D. (1997). Towards a causal model of learned hopelessness for Hong Kong adolescents. *Educational Studies*, *23*, 377-391.

Boone, W.J. (1997). Science attitudes of selected middle school students in China: A preliminary investigation of similarities and differences as a function of gender. *School Science and Mathematics*, *97*, 96-103.

Chan, D.W. (1996). Self-consciousness in Chinese college students in Hong Kong. *Personality and Individual Differences*, *21*, 557-562.

Chan, D.W. (1997). Self-concept domains and global self-worth among Chinese adolescents in Hong Kong. *Personality and Individual Differences*, *22*, 511-520.

Chan, Y.M. (2000). Self-esteem: A cross cultural comparison of British-Chinese, white British and Hong Kong Chinese children. *Education Psychology*, *20*, 59-74.

Chan, Y.M., & Verma, G.K. (1995). Self-esteem in Hong Kong pupils. *Research in Education*, *54*, 25-31.

Cheung, C.-K., & Kwok, S.-T. (1996). Conservative orientation as a determinant of hopelessness. *Journal of Social Psychology*, *136*, 333-347.

Cheung, P.C., & Lau, S. (1985). Self-esteem: its relationship to the family and school social environments among Chinese adolescents. *Youth and Society, 16*, 438-456.

Chou, K.-L. (1996). The Rushton, Chrisjohn and Fekken Self-Report Altruism Scale: A Chinese translation. *Personality and Individual Differences, 21*, 297-298.

Dorahy, M.J., & Lewis, C.A. (2001). The relationship between dissociation and religiosity: An empirical evaluation of Schumaker's theory. *Journal for the Scientific Study of Religion, 40*, 317-324.

Eckert, R.M., & Lester, D. (1997). Altruism and religiosity. *Psychological Reports, 81*, 562.

Eek, J. (2001). *Religious facilitation through intense liturgical participation: A quasi-experimental study of Swedish pilgrims to Taizé*. Lund: University of Lund Studies in Psychology of Religion.

Evans, T.E., & Francis, L.J. (1996). Measuring attitude toward Christianity through the medium of Welsh. In L.J. Francis, W.K. Kay & W.S. Campbell (eds), *Research in Religious Education* (pp 279-294). Leominster: Fowler Wright Books.

Fearn, M., Francis, L.J., & Wilcox, C. (2001). Attitude toward Christianity and psychological type: A survey among religious studies students. *Pastoral Psychology, 49*, 341-348.

Francis, L.J. (1978). Attitude and longitude: A study in measurement. *Character Potential: A record of research, 8*, 119-130.

Francis, L.J. (1987). Measuring attitudes towards Christianity among 12-18 year old pupils in Catholic schools. *Educational Research, 29*, 230-233.

Francis, L.J. (1989). Measuring attitude towards Christianity during childhood and adolescence. *Personality and Individual Differences, 10*, 695-698.

Francis, L.J. (1992). Reliability and validity of the Francis scale of attitude towards Christianity (adult). *Panorama, 4* (1), 17-19.

Francis, L.J. (1997). The psychology of gender differences in religion: A review of empirical research. *Religion, 27*, 81-96.

Francis, L.J. (1998). The relationship between intelligence and religiosity among 15 to 16 year olds. *Mental Health, Religion and Culture, 1*, 185-196.

Francis, L.J., & Greer, J.E. (1990). Measuring attitudes towards Christianity among pupils in Protestant secondary schools in Northern Ireland. *Personality and Individual Differences, 11*, 853-856.

Francis, L.J., & Greer, J.E. (1999a). Attitudes towards creationism and evolutionary theory: The debate among secondary pupils attending Catholic and Protestant schools in Northern Ireland. *Public Understanding of Science, 8*, 93-103.

Francis, L.J., & Greer, J.E. (1999b). Attitude toward Christianity among secondary pupils in Northern Ireland: Persistence of denominational differences? *British Journal of Religious Education, 21*, 175-180.

Francis, L.J., & Hermans, C.A.M. (2000). Internal consistency reliability and construct validity of the Dutch translation of the Francis scale of Attitude toward Christianity among adolescents. *Psychological Reports, 86*, 301-307.

Francis, L.J., Jones, S.H., & Wilcox, C. (1997). Religiosity and dimensions of psychological well-being among 16-19 year olds. *Journal of Christian Education, 40* (1), 15-20.

Francis, L.J., & Kwiran, M. (1999). Werthaltungen (einstellungen) gegenüber dem christentum bei deutschen heranwachsenden: Die francis-skala. *Braunschweiger Beiträge, 89* (3), 50-54.

Francis, L.J., Lewis, J.M., Philipchalk, R., Brown, L.B., & Lester, D. (1995). The internal consistency reliability and construct validity of the Francis scale of attitude towards Christianity (adult) among undergraduate students in the UK, USA, Australia and Canada. *Personality and Individual Differences, 19*, 949-953.

Francis, L.J., & McCarron, M.M. (1989). The measurement of attitudes towards Christianity among Nigerian secondary school students, *Journal of Social Psychology, 129*, 569-571.

Francis, L.J., Quesnell, M., & Lewis, C.A. (2002). Assessing attitude toward Christianity among adolescents in the Czech Republic: The Francis scale. *Ceskoslovenská Psychologie*, (in press).

Francis, L.J., & Stubbs, M.T. (1987). Measuring attitudes towards Christianity: From childhood to adulthood. *Personality and Individual Differences, 8*, 741-743.

Francis, L.J., & Wilcox, C. (1998), Religiosity and femininity: do women really hold a more positive attitude toward Christianity? *Journal for the Scientific Study of Religion, 37*, 462-469.

French, S., & Joseph, S. (1999). Religiosity and its association with happiness, purpose in life, and self-actualisation. *Mental Health, Religion and Culture, 2*, 117-120.

Fulljames, P., & Francis, L.J. (1987). The measurement of attitudes towards Christianity among Kenyan secondary school students. *Journal of Social Psychology, 127*, 407-409.

Gibson, H.M. (1989). Measuring attitudes towards Christianity among 11-16 year old pupils in non-denominational schools in Scotland. *Educational Research, 31*, 221-227.

Gibson, H.M., & Francis, L.J. (1989). Measuring attitudes towards Christianity among 11- to 16- year old pupils in Catholic schools in Scotland. *Educational Research, 31*, 65-69.

Greer, J.E., & Francis, L.J. (1991). Measuring attitudes towards Christianity among pupils in Catholic Secondary schools in Northern Ireland. *Educational Research, 33*, 70-73.

Joseph, S., & Diduca, D. (2001). Schizotypy and religiosity in 13-18 year old school pupils. *Mental Health, Religion and Culture, 4*, 63-69.

Kay, W.K., & Francis, L.J. (1996). *Drift from the churches: Attitude toward Christianity during childhood and adolescence*. Cardiff: University of Wales Press.

Kay, W.K., Francis, L.J., & Gibson, H.M. (1996). Attitude toward Christianity and the transition to formal operational thinking. *British Journal of Religious Education, 19*, 45-55.

Lau, S. (1988). The value orientations of Chinese University students in Hong Kong. *International Journal of Psychology*, *23*, 583-596.

Lau, S. (1989). Sex role orientation and domains of self-esteem. *Sex Roles*, *21*, 415-422.

Lau, S. (1990). Crisis and vulnerability in adolescent development. *Journal of Youth and Adolescence*, *19*, 111-131.

Lau, S., & Cheung, P.C. (1987), Relations between Chinese adolescents' perception of parental control and organisation and their perception of parental warmth. *Developmental Psychology*, *23*, 726-729.

Lau, S., & Leung, K. (1992a). Relations with parents and school and Chinese adolescents' self-concept, delinquency and academic performance. *British Journal of Educational Psychology*, *62*, 193-202.

Lau, S., & Leung, K. (1992b). Self-concept, delinquency, relations with parents and school and Chinese adolescents' perception of personal control. *Personality and Individual Differences*, *13*, 615-622.

Lau, S., & Wong, A.K. (1992). Value and sex-role orientation of Chinese adolescents. *International Journal of Psychology*, *27*, 3-17.

Leung, J.P. (1990). Comparison of self-concepts and delinquent behaviours in junior secondary school students from pre-vocational schools and grammar schools. *CUHK Education Journal*, *18*, 165-171.

Leung, K., & Lau, S. (1989). Effects of self-concept and perceived disapproval of delinquent behaviour in school children. *Journal of Youth and Adolescence*, *18*, 345-359.

Lewis, C.A. (1996). Religiosity and obsessionality. In L.J. Francis, W.K. Kay & W.S. Campbell (eds), *Research in Religious Education* (pp 219-227). Leominster: Gracewing.

Lewis, C.A. (1998). Towards a clarification of the association between religiosity and life satisfaction. *Journal of Beliefs and Values*, *19*, 119-122.

Lewis, C.A. (2000). The religiosity-psychoticism relationship and the two factors of social desirability: A response to Michael W. Eysenck (1999). *Mental Health, Religion and Culture*, *3*, 39-45.

Lewis, C.A., Lanigan, C., Joseph, S., & de Fockert, J. (1997). Religiosity and happiness: No evidence for an association among undergraduates. *Personality and Individual Differences*, *22*, 119-121.

Lewis, C.A., & Maltby, J. (1995). The reliability and validity of the Francis scale of attitude towards Christianity among US adults. *Psychological Reports*, *76*, 1243-1247.

Lewis, C.A., & Maltby, J. (1997). Reliability and validity of the Francis Scale of Attitude toward Christianity (adult) among Northern Irish university students. *Irish Journal of Psychology*, *18*, 349-354.

Lewis, C.A., & Maltby, J. (2000). Conservatism and attitude toward Christianity. *Personality and Individual Differences*, *29*, 793-798.

Livingston, S.A. (1988). Reliability of test results. In J.P. Keeves (ed.), *Educational Research, Methodology and Measurement: An international handbook* (pp 386-392). Oxford: Pergamon Press.

Lu, L., & Shih, J.B. (1997). Personality and happiness: is mental health a mediator? *Personality and Individual Differences*, *22*, 249-256.

Lu, L., Shih, J.B., Lin, Y.Y., & Ju, L.S. (1997). Personal and environmental correlates of happiness. *Personality and Individual Differences, 23,* 453-462.

Maltby, J. (1994). The reliability and validity of the Francis scale of attitude towards Christianity among Republic of Ireland adults. *Irish Journal of Psychology, 15,* 595-598.

Man, K., & Hamid, P.N. (1998). The relationship between attachment prototypes, self-esteem, loneliness, and causal attributions in Chinese trainee teachers. *Personality and Individual Differences, 24,* 357-371.

Munayer, S.J. (2000). The ethic identity of Palestinian Arab Christian adolescents in Israel. Unpublished PhD dissertation, University of Wales (Oxford Centre for Mission Studies).

Ng, H-S., Cooper, M., & Chandler, P. (1998). The Eysenckian personality structure: A 'Giant Three' a 'Big Five' model in Hong Kong. *Personality and Individual Differences, 25,* 1111-1131.

Orton, R.E. (1987). The foundations of construct validity: Towards an update. *Journal of Research and Development in Education, 21,* 22-35.

Schludermann, E.H., Schludermann, S.M., & Huynh, C.-L., (2000). Religiosity, prosocial values, and adjustment among students in Catholic high schools in Canada. *Journal of Beliefs and Values, 21,* 99-115.

Stigler, J.W., Smith, S., & Mao, L. (1985). The self-perception of competence by Chinese children. *Child Development, 56,* 1259-1270.

Youtika, A., Joseph, S., & Diduca, D. (1999), Personality and religiosity in a Greek Christian Orthodox sample. *Mental Health, Religion and Culture, 2,* 71-74.

Zeller, R.A. (1988). Validity. In J.P. Keeves (ed.), *Educational Research, Methodology, and Measurement: An international handbook* (pp 322-330). Oxford: Pergamon Press.

Young Minority Females' Perceptions of Sexual Abuse: A Focus Group Approach

Ruth S. Buzi
Baylor College of Medicine

Susan R. Tortolero
University of Texas Health Science Center at Houston

Peggy B. Smith
Baylor College of Medicine

Michael W. Ross
University of Texas

Robert E. Roberts
University of Texas Health Science Center at Houston

The purpose of this study was to examine female adolescents' perceptions about issues related to a sexual abuse experience. Participants were 60 African-American and Hispanic female adolescents receiving family planning services at a Teen Health Clinic. Content analysis was used to identify and organize patterns in the focus group data. Participants perceived sexual abuse to be a significant problem in the community. They based their perceptions on personal experiences, knowing friends who have been sexually abused, and information gathered from the media. Participants perceived sexual abuse to be a concept that incorporates various contact and non-contact behaviors that are nonconsensual and violate the victim. They identified a wide range of psychological effects that result from such an experience including sexual risk taking behaviors. They also validated concerns that the prevalence of sexual abuse is underestimated due to underreporting. Participants identified primarily family members, family friends, fathers, and surrogate fathers as perpetrators. Adolescents identified individual, family, community, and cultural factors that promote sexual abuse. The complexity of sexual abuse emphasizes the need for multilevel interventions focusing on the societal forces that contribute to its occurrence.

Author info: Correspondence should be sent to: Dr. Ruth Buzi, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030.

North American Journal of Psychology, 2002, Vol. 4, No. 3, 441-456.
© NAJP

It is estimated that 7% to 17% of American teenage girls have been victims of a coercive sexual experience during childhood or adolescence (Erickson & Rapkin, 1991; Nagy, Adcock, & Nagy, 1994; Nagy, DiClemente, & Adcock, 1995). The experience of sexual abuse during childhood and adolescence has been associated with a broad spectrum of poor mental health outcomes for females as well as males. It has been linked to depression (Erickson & Rapkin, 1991), suicidal ideation or attempts (Nagy, DiClemente, & Adcock, 1995), antisocial behavior (Nelson, Higginson & Grant-Worley, 1994), and frequent alcohol and other substance use (Wilsnack, Vogeltanz, Klassen, & Harris, 1997).

Recently, the role of a history of sexual abuse as a factor in risky sexual behaviors has been examined. For example, the 1997 Massachusetts Youth Risk Behavior Surveillance Survey (YRBSS) assessed whether adolescents with a history of sexual abuse were more likely than adolescents without such a history to engage in sexual risk taking behavior (Raj, Silverman, & Amaro, 2000). This study included 1,610 sexually experienced girls and 831 sexually experienced boys. The results suggested that 30.2% of girls and 9.3% of boys disclosed a sexual contact against their will. Sexually abused girls were significantly more likely than those without such a history to engage in early sexual activity (Odds Ratio=2.2, 95% Confidence Interval =1.46-3.47), to report multiple partners (Odds Ratio= 1.9, 95% Confidence Interval= 1.21-2.92) and to have been pregnant (Odds Ratio=3.4, 95% Confidence Interval= 1.53-7.34).

Fergusson, Horwood, and Lynsky (1997) examined the extent to which exposure to childhood sexual abuse was associated with increased rates of sexual risk taking behaviors. At age 18, retrospective reports of childhood sexual abuse were obtained from participants. The results suggested that females who experienced childhood sexual abuse with intercourse were 4.4 times more likely to engage in sexual activity and be pregnant, 9.3 times more likely to have multiple sexual partners, and 6.6 times more likely to have their first sexual intercourse before age 16, than non-abused females. The association remained significant even after controlling for other adverse childhood experiences.

Cohen et al., (2000) also examined the effects of childhood sexual abuse in women with HIV or at risk for HIV infection. They studied 1,288 women with HIV and 357 uninfected women in order to determine the prevalence and effect of domestic violence and childhood abuse on sexual behaviors. The results of the study suggested that 31% of the HIV-seropositive women and 27% of the HIV-seronegative women reported childhood sexual abuse. Childhood sexual abuse was strongly associated with having more than 10 male sexual partners, having male partners at risk for HIV infection, and exchanging sex for drugs, money, or shelter.

Although numerous studies have suggested that sexual abuse is prevalent and that a history of sexual abuse produces a cluster of negative effects, including sexual risk taking behaviors, these studies have multiple methodological shortcomings. In a recent review, Goldman and Padayachi (2000) suggested that disparate definitions of child sexual abuse, number of screening questions, and choice of a sample are some of the methodological problems. Others have suggested that in order to understand fully the experience of sexual abuse, it is essential to learn the meaning and definition of this experience for women, taking into consideration their cultural background and social context (Ramos Lira, Koss, & Felipe Russo, 1999). Although studies that examined aspects of sexual abuse issues among adolescents provide valuable information, these studies do not integrate information about adolescent females' perceptions on these issues. It is suggested that research approaches that are attentive to the experiential world of adolescent females can assist in understanding the specific concerns of females in this developmental stage (Banister & Schreiber, 2001).

The focus group methodology is a useful research tool that is designed to determine how people regard an experience or event through understanding their personal perceptions and interpretations (Krueger, 1998). However, focus groups appear to be underutilized in sexual abuse research. It is suggested that this is a result of the common perception that the methodology requires sensitivity and that there is a lack of understanding in the field regarding the use of focus groups (Charlesworth & Rodwell, 1997). The purpose of the present study was to provide young females the opportunity to share their perceptions and concerns about sexual abuse issues in the community. This information can be valuable in understanding the impact of a sexual abuse experience on young females and can help in designing interventions for this age group.

METHODS

Sample

Females ranging in age from 16-22 who attended the Teen Health Clinic between September 2001 to February 2002 were recruited to participate in the study. The Teen Health Clinic is located in a large county hospital in the southwest part of the United States. This clinic provides free of charge comprehensive family planning and reproductive health services to indigent adolescents who reside in the inner city. Parental consent for clinical services is solicited but not required because minors who receive services are not required to have parental consent by state statutes.

Participants were recruited by the lead author of this paper and a staff social worker. Participants who met the study criteria were approached on designated days and asked to participate in the focus groups. The focus group meetings were conducted in a private room adjacent to the waiting room. Some of the subjects participated in the focus groups while waiting for clinical services, while others were invited to participate and came to the clinic on the day of the focus group specifically for this purpose. The purpose of the study was explained and informed consent was obtained prior to group discussions. Participation was voluntary and participants were asked permission to audio-tape group discussions. In order to ensure participants' confidentiality, participants were given tags with different names. The study protocol was approved by the Institutional Review Board of the affiliated school.

Measures

The following questions were presented to adolescents participating in the focus groups: What is sexual abuse? Is sexual abuse a problem in the community? What are the consequences for children who have been sexually abused as they grow up? Are adolescents who have been sexually abused more likely to engage in sexual risk taking behavior? If it is related, how can the association be explained? What kinds of services are needed by adolescents who have been abused? What can the clinic do to help adolescents?

Procedures

The group process was driven by the above five questions. Since sexual abuse can be a sensitive and stigmatizing phenomenon, the questions were not designed to elicit information on personal experience with abuse. At the beginning of each focus group, the moderator explained that the purpose of the focus group was to obtain from participants input about their perceptions of sexual abuse in the community and its consequences for the victims. It was emphasized that participants were not expected to disclose personal information. The moderator explained issues related to confidentiality and clarified that if personal information about abuse for minors was presented, appropriate investigation and notification of authorities might be necessary. Data were audio taped and the moderators also took notes. An instrument recording basic demographic information such as age, ethnicity, and school status was administered prior to each session. The questions for this paper were presented to participants as part of a larger study designed to elicit information on issues related to resiliency and protective factors. Due to the sensitive nature of the topic, these questions were presented toward the end of the meeting after rapport was established. At the end of each session participants were encouraged to discuss any issue or ideas with the moderators without being audio taped. Groups consisted of 3-7 participants. Each group met for one session, lasting 75-90 minutes. A total of 11 groups were conducted. Participants received a $ 20 gift certificate for their time.

Data Analysis

Transcripts of the discussions were prepared by a research assistant and included every statement made in the group by each participant. Transcripts were imported into the Qualitative Solutions and Research (QSR) NUD*IST Vivo (NVivo), version 1.1, a computerized software package that aids in the management and analysis of qualitative data (Qualitative Solutions & Research Pty, Ltd, 1999). Nodes and attributes, representing defined categories and ideas, were created initially and tracked throughout the group sessions. Additional nodes and attributes were created throughout the process when new ideas were identified. Emerging themes and concepts were identified and organized. Recurring themes were also identified by comparing the transcribed data to the notes taken by the moderators.

RESULTS

Participants' mean age was 18.13 (range: 16-22; SD: 1.65). Eighty percent were African American and 20% were Hispanic. Fifty-five percent were in school and 45% were not in school.

Eight major themes emerged from the data. These themes, some of which related to the specific questions posed at the focus groups, are presented with selected quotes to illustrate them.

Definition of Abuse

Participants in the focus groups defined sexual abuse as a broad and inclusive term encompassing a wide range of behaviors. They construed sexual abuse as a continuum, spanning sexual kissing, fondling, touching, and exhibitionism, oral sex, genital, and anal sex. The majority of participants defined sexual abuse as any type of act that is not consensual, which makes the victim feel uncomfortable and/or violated. The definition considered the behaviors involved, the non-consent of victims, and the subjective feelings of the victims. There was some disagreement as to whether experiencing sexually suggestive remarks by men is also considered sexual abuse. Some participants considered sexual abuse only as forced sexual intercourse. Participants did not limit the age at the time of the sexual abuse experience or the age of the perpetrator, for the experience to be considered sexual abuse.

Perception of sexual abuse as a problem in the community

There was ample evidence in these focus groups that inner-city females were concerned about sexual abuse in their community. In every group, some participants said they personally knew friends or acquaintances who were victims of sexual abuse. In the initial discussions, participants spoke only about experiences of other females they knew; after a while, some also disclosed their own experiences. Participants also used information they gathered from the media about cases of sexual abuse of children to support their belief that sexual abuse is a significant problem in the community. A group of college freshmen talked about the pervasiveness of sexual harassment and degrading attitudes of males on campus including sexual remarks and suggestions, inappropriate touching, and coercion for having sex when under the influence of alcohol. Some of the responses to the question about the sources of the knowledge were as follows: "You can see it with your eyes," "You can tell how it is," " I know some of them," It happened to me," "It happened to my friend," "It happened to my sister."

Disclosure of a sexual abuse experience

Focus group participants suggested that many victims of sexual abuse are reluctant to disclose the sexual abuse experience. Fear, shame and embarrassment were identified as the main reasons for the silence of victims. Victims do not disclose it because of fear of being perceived as damaged or tarnished by the family or friends. Participants also indicated that it is hard to find someone who can be trusted and who will not reveal the abuse to others. Participants repeatedly indicated that there are many children and adolescents who never disclose these experiences. It was not clear if they referred to their own experiences. One of the responses that this question generated was as follows:

"Even though you have like a girlfriend, some stuff you just don't want to tell her. You don't want to talk about it. You don't know how they will see you. Sometimes people will talk to you different just by what they know, or sometimes you don't want them to know that. You want them to see you as one person. You don't want them to see the faults in you."

Participants emphasized the complexity of disclosure when the perpetrator is a close family member. The fear and uncertainty of the family's reaction to the disclosure silence victims. These concerns also indicated fear of the financial consequences of such a disclosure to the family. One of the participants suggested:

"How can you tell your parent if your parent is involved? Say it's your mom's friend and he brings your mother money. How can you tell your mom?"

Attribution of blame

In the focus groups, perpetrator blame was common. Regarding perpetrators, participants said, "These persons are sick, something is wrong with them, and these guys aren't mentally stable." Talking about these perpetrators provoked strong emotional responses of anger and disgust. There were also suggestions that many of these perpetrators were molested as children and they repeat the cycle of abuse. Although perpetrator blame was common, participants also attributed blame to parents who fail to provide children a protective environment. They indicated that often the perpetrator is the stepfather or the mother's boyfriend. The mother denies the accusations of the child because she is financially dependent on these men. They repeatedly talked about mothers who preferred men over their children. Participants described strong feelings of betrayal and abandonment associated with the mother's disbelief and denial of the accusations. Participants also described situations in which the mother was a drug addict and depended on men to support her habits. These men can victimize a child in the family while the mother does not protect the child from risks.

Some of the responses to parents' reactions were the following:

"It was my mamma's boyfriend and you know there are a lot of women that believe their boyfriend before they believe their child."

"But if you try to tell your parents and your parents reject you, that's very painful, because every time you think of your parents, you try...everybody think of the good things their parents have done for them. But now if you go back and think of something like this, my parents don't believe me, it hurts because your mamma and your daddy are supposed to stand by you."

There were also some attributions of self-blame in some groups. A few participants suggested that many young girls actually consented to having sex with older males in a search for a father figure.

Perceived Consequences of a Sexual Abuse Experience

Participants identified a wide range of painful psychological effects that result from a sexual abuse experience, including low self-esteem, depression, drug use, helplessness, suicide, and re-victimization. Participants felt that such experiences had a long-lasting effect. In regard to sexual behaviors, participants suggested that the majority of victims will become promiscuous and will engage in frequent sex with multiple partners. They also indicated that teenage pregnancy is many times associated with a sexual abuse experience at a young age. However, some suggested that victims also can isolate themselves as a result of the experience and have difficulties trusting men. When asked if the type of psychological symptoms experienced depends on the type of abuse experienced, consensus among participants was that any type of act will result in distress among victims. A few felt that in cases of intercourse, the results for the victims are more severe than in non-contact forms of abuse. Some of the responses to this question were as follows:

"They try to ease their pain. They might start drinking or doing drugs, getting into a lot of problems in school, getting into fights. Just their whole outlook on life changed."

"They take it out on smoking, drugs. They take it out on having sex with other people to get it off their minds."

Gender Issues

Participants identified predominantly adult males as perpetrators. These adults were primarily father surrogates, family members, family friends, acquaintances, and fathers. A few participants also identified teachers and church clergy as perpetrators. They indicated that females could be the perpetrators but that men abuse children more frequently. Participants also indicated that females are more likely to become victims of sexual abuse but felt that boys are victims as well. The question whether females and males perceive the sexual abuse similarly provoked some disagreement. Many participants felt that a sexual abuse experience is not as traumatic for males as it is for females. They indicated that many older boys perceive that experience, especially when the perpetrator is a female, as sex education or sex initiation. However, some participants felt that because sexual abuse of females is more common and because males are expected to defend themselves, being a victim is a more difficult experience for males. Several participants suggested that one of the consequences for males who were abused by other males was homosexuality, a consequence that was not often mentioned for females. They indicated that males were even less likely than females to disclose the sexual abuse experience if they were victimized by males, because of the stigma associated with it.

One of the comments to the issue of male abuse was as follows:

"I think it's more painful for the boys because it's more common to happen to girls. If it happens to a boy, a boy is expected to protect himself and be a man."

Females felt that in general the world is an unsafe place and that females are prone to a variety of sexual abuse forms by males. Participants expressed negative views and mistrust towards males. Referring to males, one of the participants said:

"A dog is a dog. You (are) a stupid one or smart one, they're all dogs. I think all men are no good, it just you gotta pick which one is better."

Mechanism explaining subsequent problems

Participants offered explanations as to how a sexual abuse experience leads to behavior problems. They felt that the sexual abuse experience is a traumatic event. Consequently, victims of sexual abuse self-medicate with drugs and alcohol in order to maintain the dissociation of painful memories. They had several explanations as to how promiscuity results from a sexual abuse experience. One of the explanations was that the experience distorts the victim's perception of sex, which becomes acceptable and normative. Another explanation was that it damages the victim's self-esteem and as a result she does not have respect for herself and engages in sex with multiple partners. They also indicated that the victims are seeking attention and love through sexual activity. Participants also suggested that the sexual abuse experience seriously damages the perception of sexual activity as a romantic affair. These ideas were reflected by the following responses:

"They feel like they're not worth anything. They are soiled and dirty, so they might as well behave dirty."

"Sex is supposed to be about love, relationships, you are with a person you are in love with. But when someone violates you, it makes you feel like sex is just nothing, sex is dirty. You are giving it up to anybody cause it has no meaning anymore."

Prevention of sexual abuse

Participants identified several areas that need to be targeted for prevention. They suggested that it is important to reach out to children and adolescents and provide victims of sexual abuse safe places where they can disclose a sexual abuse experience. They suggested that individual as well as group counseling is essential to help teens deal with the sexual abuse experience. Counseling can be provided at clinics or schools. Participants emphasized the importance of privacy and confidentiality in providing counseling. They also suggested that in order to prevent sexual abuse, sex education classes in schools have to teach children and adolescents how to differentiate between healthy and unhealthy relationships. They emphasized that a preaching style is not effective but suggested instead engaging children and adolescents in discussions to help them clarify issues related to healthy sexuality. This was reflected by the following response of one of the participants:

"Children have a lot of misconceptions about what sex is. You need to ask them. You need feedback. You need to know where they are coming from. Then when they tell you sex is when you touch somebody and they may not know, then you explain to them and clarify any question that they have. Then show them stuff that could happen; then they have a reality of what could happen."

Suggestions were also made to encourage and educate parents to address the root causes of inadequate nurturing, supervision, and monitoring of children. Participants also suggested that parents need to be targeted for interventions that will teach them how to communicate the risk of being abused to children.

DISCUSSION

The findings from these focus groups clearly indicate that adolescents perceive sexual abuse to be a problem in the community that carries major consequences for the victims. These adolescents also validated concern that sexual abuse is underestimated due to underreporting. The ambiguity regarding the definition of the term sexual abuse in these groups reflects the debate about this issue in the literature. The term sexual abuse has never been clearly defined in the literature (Hibbs, 1998). The lack of a commonly accepted definition of the term sexual abuse hinders research, treatment, and prevention efforts (Haugaard, 2000). Definitional problems also limit the ability to document accurately the prevalence of sexual abuse because rates vary based on the definition used in the study (Hibbs, 1998). The majority of participants in this study viewed sexual abuse as a broad concept involving non-contact as well as contact sexual experiences. They all agreed that non-consensuality is what makes the act abusive regardless of the type of the experience. This suggests that studies should use a more detailed definition of abuse when querying about sexual abuse experiences and that non-consensuality needs to be part of the definition. This also supports assertions by some researchers that health professionals need to expand their definition of narrowly defined ideas of unwanted sexual experiences to include more subtle forms of unwanted sexual experiences. It is suggested that studies focus mainly on the sequelae of the more severe forms of sexual abuse such as childhood sexual abuse, rape, and incest, while ignoring more subtle forms of unwanted sexual experiences (Erickson & Rapkin, 1991).

According to the National Child Abuse and Neglect Data System (NCANDS), cases of substantiated sexual abuse have declined approximately 39% nationwide from 1992 to 1999. Some of the possible causes for the decline include increased evidentiary requirements to substantiate cases, increased caution of the investigators due to new legal rights of caregivers, and a real decline in occurrence (Jones, Finkelhor & Kopiec, 2001). Despite a reported decline of sexual abuse cases, females in this study perceived sexual abuse to be a significant problem in their community.

Disclosure of sexual abuse is critical in initiating interventions to end the abuse, prosecuting the perpetrator, and providing therapeutic services to the victim to decrease the likelihood of long-term negative outcomes (Paine & Hansen, 2002). Reports of adolescents in these focus groups validate assertions that the actual prevalence of sexual abuse is underestimated due to underreporting (Finkelhor, Hotaling, Lewis, & Smith, 1990). Adolescents have identified a diversity of socio-cultural factors such as attitudes toward purity or stigmatization of those sexually abused that impede self-disclosure of sexual abuse. One of most critical obstacles to disclosure that participants identified was the failure of family members to believe or support the disclosure. Participants' concerns about possible negative reaction associated with disclosure makes sense in light of the findings that telling victims they are to blame for the assault or were irresponsible following disclosure were strongly related to increased psychological symptomatology and delayed recovery (Campbell, Ahrens, Sefl, Wasco, & Barnes, 2001).

Attributions of responsibility and blame for the sexual abuse experience appeared to be issues that concerned participants. While the attributions of responsibility for the sexual abuse experience can be directed at oneself, the perpetrator, and the family, self-blame attributions are common and have been a focus in the literature (McMillen & Zuravin, 1997). Self-blame attributions can result from victims feeling that they participated in the sexual behavior, failed to disclose the abuse, failed to stop the abuse, failed to protect siblings, and gained pleasure (Celano, 1992). While the focus in the literature has been self-blame (McMillen & Zuravin, 1997), adolescents in these focus groups attributed blame to the perpetrators and the family. Attributions of responsibility are important because they have implications for how one views the world and other, and may explain the link between childhood sexual abuse and relationships problems (Barker-Collo, 2001; McMillen & Zuravin, 1997). Therefore, it is important to assess these attributions when working with sexual abuse victims.

Participants in this study had a very good understanding about the myriad poor mental health outcomes associated with this experience. These perceptions support findings of studies that suggest women's experience of sexual abuse in childhood and adolescence may be an important risk factor for later psychopathology (Brown, Lourie, Zlotnick, & Cohn, 2000; Fergusson, Horwood, & Lynsky 1997), especially inappropriate sexual behaviors (Fergusson, Horwood, & Lynsky, 1997; Raj, Silverman, & Amaro, 2000).

The discussions in these focus groups also reflected the pervasiveness of sexual violence against women and girls in American society, especially by the men closest to them. In order to ameliorate violence against females, it is important to understand societal factors that lead to violent behaviors. It is suggested that the process of socializing boys into men in this culture promotes aggression and toughness. The dominant view of masculinity identity portrays men as tough, aggressive, and sexually active while females are socialized to have an inferior role in society (Crocco, 2001). Adolescents in focus groups also verbalized strong feelings of homophobia when they identified homosexuality as a consequence of a sexual abuse experience. It is suggested that homophobia is related to the same cultural views that idolize sexualized male supremacy and degrade behaviors that are associated with females (Crocco, 2001). The information provided by participants lends support to the need to address gender role stereotypes that are pervasive in our society. In order to change these gender role stereotypes, boys need alternative images of masculinity. Schools are suggested as sites where curricula can be implemented which help children understand and resist gender violence (Crocco, 2001).

Adolescents' explanations regarding the mechanism that explains how a sexual abuse experience leads to maladaptive behaviors especially in regard to sexuality are consistent with some of the theories that have been suggested to explain this association. For example, Finkelhor and Browne (1985) proposed a model that describes four trauma-causing factors: traumatic sexualization, betrayal, stigmatization, and powerlessness that interrupt normal development. Bandura (1977) suggested that a sexual abuse experience may lead to learned maladaptive behaviors, beliefs, and attitude through modeling and reinforcement by the perpetrator. In addition to these explanations, adolescents had some ideas about the "death" of romanticism that is caused by the sexual abuse experience that leads to sexual risk behaviors.

Adolescents' perceptions about the definition of sexual abuse, attribution of blame, and nondisclosure are similar to the views reported by unacculturated Latina women (Ramos Lira, Koss & Felipe Russo, 1999). This may suggest that ethnicity or acculturation are not as significant as socioeconomic and gender relations in shaping females' experiences. Indeed, sexual assault against women is widespread around the world (Koss, Heise & Felipe Russo, 1994). The pervasiveness of violence against women raises questions about broader societal forces that promote it.

Adolescents in focus groups identified individual, family, community, and cultural factors that support the occurrence of child sexual abuse. This included factors such as individual pathology, poverty, dysfunctional family structure, substance abuse, and gender relations that promote sexual abuse. They identified poverty as a factor that locks families in compromising relationships that have severe consequences for children. These observations support the adoption of an ecological model that explains maltreatment of children. Bronfenbrenner's (1979) ecological systems theory has been expanded by Sidebotham and Golding (2001) to reflect how child maltreatment is determined by interactions of multiple forces at four levels. The first level is ontogenic development, the processes that determine how a parent becomes abusive or neglectful. The second level is the microsystem, the context of the child's immediate environment. The third level is the exosystem, the larger societal system that surrounds the family. Finally, there is the macrosystem, the overarching cultural beliefs and norms in any particular society. The identification of multiple factors that contribute to sexual abuse by focus group participants and the support from the literature suggests that preventing sexual abuse requires multilevel interventions.

Several important implications for prevention and research may be drawn from the discussions with these adolescents. First, incorporating unwanted sexual experiences awareness into curricula on adolescent sexuality is essential. These classes must address the broader issues involved in sexuality such as building healthy relationships. Second, conducting routine screening for past or current sexual abuse experiences, including the more subtle forms, and providing referrals was also supported by these adolescents. Third, interventions should focus on how to avoid negative reactions when attempting to provide support to victims of sexual abuse. Finally, further research that focuses on understanding societal and cultural factors that promote violence against women is crucial.

These results cannot be taken as generally representing the perceptions of adolescents from other population groups. Since the groups were ethnically mixed, possibly different ethnic nuances might have been diluted. There also was no representation of majority group females. For future studies, structuring groups by ethnicity and ensuring a better representation of ethnic groups could help identify possible ethnic differences more accurately. Despite these limitations, this study provides useful data on adolescents' views of sexual abuse issues. Since sexual abuse is a complex issue, it is important that researchers provide opportunities for individuals to express their perceptions.

Although focus groups appear to be an important tool in understanding the meaning of unwanted sexual experiences among adolescents, there are some ethical issues that need to be considered due to the sensitive nature of the topic. The focus group dynamic is similar to that of therapy groups, where the synergistic effect that is developed in the group process can lead to over-disclosure of personal information (Smith, 1995). This over-disclosure might entail disclosure of information that requires further investigation and possible reporting to the proper authorities. For example, despite notifying participants about the limits of confidentiality, a 17-year-old participant disclosed a personal sexual abuse experience that required further investigation. Smith (1995) reviews specific ethical issues in focus group research and suggests several recommendations that need to be followed in order to ensure compliance with the ethical principles of autonomy, beneficence, nonmaleficence, and justice in working with focus groups. This includes informing participants that confidentiality is limited, forming groups with fewer participants to allow better group cohesion, and to employ two moderators to facilitate better control of group dynamics. It is also important to employ moderators with clinical skills who will know how to deal with disclosed sensitive information. With careful planning and selection of subjects, the focus group can provide valuable insight for the study of sexual abuse.

REFERENCES

Bandura, A. (1977). Social Learning Theory. Englewood Cliffs, NJ: Prentice-Hall.

Banister, E., & Schreiber, R. (2001). Young women's health concerns: Revealing paradox. *Health Care for Women International, 22*, 633-647.

Barker-Collo, S. L., 2001. Adult reports of child and adult attributions of blame for childhood sexual abuse: predicting adult adjustment and suicidal behaviors in females. *Child Abuse & Neglect, 25* (10), 1329-1341.

Bronfenbrenner, U. (1979). *The ecology of human development*. Cambridge, MA: Harvard University Press.

Brown, L. K., Lourie, K. J., Zlotnick, C., & Cohn, J. (2000). Impact of sexual abuse on the HIV-risk-related behavior of adolescents in intensive psychiatric treatment. *American Journal of Psychiatry, 157* (9), 1413-1415.

Campbell, R., Ahrens, C. E., Sefl, T., Wasco, S. M., & Barnes, H. E. (2001). Social reactions to rape victims: healing and hurtful effects on psychological and physical health outcomes. *Violence & Victims, 16* (3), 287-302.

Celano, M. P.(1992). A developmental model of victims' infernal attributions of responsibility for sexual abuse. *Journal of Interpersonal Violence, 7*, 57-69.

Charlesworth, L.W. & Rodwell, M.K. (1997). Focus groups with children: a resource for sexual abuse prevention program evaluation. *Child Abuse & Neglect, 21* (12), 1205-1206.

Cohen, M, Deamant, C., Barkan, S., Richardson, J., Young, M., Holman, S., et al. (2000). Domestic violence and childhood sexual abuse in HIV-infected women and women at risk for HIV. *American Journal of Public Health, 90* (4), 560-565.

Crocco, M. S., (2001). The missing discourse about gender and sexuality in social studies. *Theory Into Practice, 40* (1), 65-72.

Erickson, P. I. & Rapkin, A. J. (1991). Unwanted sexual experiences among middle and high school youth. *Journal of Adolescent Health, 12*, 319-325.

Fergusson, D. M., Horwood, J., & Lynsky, M. T. (1997). Childhood sexual abuse, adolescent sexual behaviors and sexual revictimization. *Child Abuse & Neglect, 21* (8), 789-803.

Finkelhor, D. & Browne A. (1985). The traumatic impact of child sexual abuse: a conceptualization. *American Journal of Orthopsychiatry*, 55 (4), 530-541.

Finkelhor, D., Hotaling, G., Lewis, I. A., & Smith, C. (1990). Sexual abuse in a national survey of adult men and women: prevalence, characteristics, and risk factors. *Child Abuse & Neglect, 14* (1),19-28.

Goldman, J.D.G. & Padayachi, U. K. (2000). Some methodological problems in estimating incidence and prevalence in child sexual abuse research. *The Journal of Sex Research, 37* (4), 305-314.

Haugaard, J. J. (2000). The challenge of defining child sexual abuse. *American Psychologist, 55* (9), 1036-1039.

Hibbs, E. D. (1998). Improving methodologies for the treatment of a child and adolescent disorders: Introduction. *Journal of Abnormal Child Psychology, 26* (1), 1-6.

Jones, L. M., Finkelhor, D., & Kopiec, K. (2001). Why is sexual abuse declining? A survey of state child protection administrators. *Child Abuse & Neglect, 25*(9), 1139-1158.

Koss, M. P., Heise, L., & Felipe Russo, 1994). The global health burden of rape. *Psychology of Women Quarterly, 18*, 509-537.

Krueger, R. A. (1998). *Analyzing and reporting focus groups results*. Thousand Oaks, CA: Sage Publications.

McMillen, C & Zuravin, S. (1997). Attributions of blame and responsibility for child sexual abuse and adult adjustment. *Journal of Interpersonal Violence, 12* (1), 30-49.

Nagy, S., Adcock, A. G., & Nagy, M. C. (1994). A comparison of risky health behaviors of sexually active, sexually abused, and abstaining adolescents. *Pediatrics, 93*, 570-575.

Nagy, S., DiClemente, R., & Adcock, A. G. (1995). Adverse factors associated with forced sex among southern adolescent girls. *Pediatrics, 96*, 944-946.

Nelson, D. D., Higginson, G.K., Grant-Worley, J.A. (1994). Using the Youth Risk Behavior Survey to estimate the prevalence of sexual abuse among Oregon high school students. Journal of School Health, *64*, 413-416.

Paine, M. P. & Hansen, D. J. (2002). Factors influencing children to self-disclose sexual abuse. *Clinical Psychology Review, 22*, 271-295.

Raj, A., Silverman, J. G., & Amro, H. (2000). The relationship between sexual abuse and sexual risk among high school students: Findings from the 1997 Massachusetts Youth Risk Behavior Survey. *Maternal and Child Health Journal, 4* (2), 125-134.

Ramos Lira, L., Koss, M. P., & Felipe Russo, N. (1999). Mexican American women's definitions of rape and sexual abuse. *Hispanic Journal of Behavioral Sciences, 21* (3), 236-265.

Sidebotham, P. & Golding, J. (2001). Child maltreatment in the "children of the nineties" a longitudinal study of parental risk factors. *Child Abuse & Neglect, 25* (9), 1177-1200.

Smith, M. W. (1995). Ethics in focus groups: A few concerns. *Qualitative Health Research, 5*, 478-487.

Wilsnack, S. C., Vogeltanz, N. D., Klassen, A. D., & Harris, T.R. (1997). Childhood sexual abuse and women's substance abuse: national survey findings: Journal of Studies of Alcohol, *58* (3), 264-71.

Mood, Rumination, and Mood Awareness in the Athletic Performance of Collegiate Tennis Players

Vann B. Scott, Jr., Kamee B. Stiles,
David B. Raines, & Andreas W. Koth
Armstrong Atlantic State University

The relationships between athletic performance and several key variables, such as mood, rumination, and mood awareness were explored in a sample of college tennis players. Both mood awareness and rumination were negatively correlated with athletic performance. An individual differences approach to understanding the relationships among these variables, particularly with regard to mood awareness, is advocated and the implications of this approach for alleviation of the deleterious effects of negative moods on athletic performance are discussed.

For at least the past thirty years, researchers have been attempting to determine the variables that best differentiate athletic performance. Some success in differentiating athletic performance has been demonstrated using such variables as anxiety, stress, and mood (e.g., Anshel & Wells, 2000; Gould & Urdy, 1994; Terry, Coakley, & Karageorghis, 1995). The relationship between mood and athletic performance, however, remains a bit of an enigma. Despite being one of the most highly explored variables in this area of research with more than 250 studies having been conducted to explore the relationship between mood and sport/exercise performance, little has been learned regarding the true nature of the relationship between these variables (Beedie, Terry, & Lane, 2000). Some researchers have had success in demonstrating the relationship between mood and athletic performance. Newby and Simpson (1994) found that mood related to final season statistics in college volleyball players, while Friend and LeUnes (1990) differentiated baseball player performance using a mood measure. Other researchers, however, are skeptical of the usefulness of mood as a way to differentiate athletic performance (Rowley, Landers, Kyllo, & Ethier, 1995; Prapavessis, 2000).

In his review of the literature on the use of the POMS as a measure of mood to differentiate athletic performance, Prapavessis (2000) concluded that the ability of researchers to differentiate levels of athletic

---

Author info: Correspondence should be sent to: Vann B. Scott, Jr., Ph.D. Department of Psychology, Armstrong Atlantic State University 11935 Abercorn Street, Savannah, GA 31419

North American Journal of Psychology, 2002, Vol. 4, No. 3, 457-468.

© NAJP

performance from mood is moderated by the skill level of the athlete. Particularly within the group of elite athletes, the POMS may not differentiate athletic performance, in part because of the iceberg profile – a tendency for elite athletes to demonstrate a predominantly positive mood profile. Other researchers agree that the iceberg profile can make the use of the POMS for measuring mood in elite athletes problematic and advocate cautious interpretation of results using the POMS. Renger (1993) reviewed the POMS as a mood measure in sport settings and concluded that the POMS serves to differentiate the performance of the athlete from the non-athlete but has shown less success at differentiating the successful athlete from the less successful athlete. This position is also supported by the work of Beedie et al. (2000).

Due to the equivocal nature of the findings of studies exploring the power of mood (as measured by the POMS) to differentiate athletic performance, the present study investigates the usefulness of alternative variables for differentiating athletic performance. The specific variables of interest are mood, rumination (worry), and mood awareness. These three variables are hypothesized to be related to the performance of the athlete via a relationship to cognitive processing resources. The rationale for the use of these variables will be elaborated upon in separate, variable-specific sections below.

Watson, Clark, and Tellegen (1988) have developed a measure of mood (the Positive Affect Negative Affect Schedule; PANAS) that conceptualizes it as both a trait characteristic and a temporary state. Mood, as measured by the PANAS, provides a versatile tool for determining the impact of mood traits and mood states. In the past, researchers exploring the differentiation of athletic performance with mood measures often conceptualized mood only in the typical state form. However, much research has focused on the concept of negative affectivity (or neuroticism) and its tendency to taint the processing of information (Watson et al., 1988). The trait form of mood is important to the current study as it is expected to have a more enduring effect on the cognitive processing of information. People who experience greater levels of negative affect tend to view the world as more harsh and less satisfying and will tend to see relatively neutral comments directed toward them as negative attacks (Watson et al., 1988). The propensity toward negative affectivity also is associated with increased levels of dysphoria, depression, and perseveration on the negative mood state, which in turn consumes cognitive resources. Watson et al. (1988) found that the PANAS has acceptable reliability and validity indicators in a variety of samples including college students, university employees, and psychiatric patients. In a sample of 70 high school wrestlers, the PANAS was found to have similar internal consistency values (.84 for PA and .79

for NA) to those reported by Watson et al. 1988 (Treasure, Monson, & Lox, 1996). A heterogeneous sample of 66 rugby players (age in years, $M = 26.8$, $SD = 6.3$) showed similar internal consistency values for the PA and NA scales as well (.86 & .71, respectively).

Swinkels and Giuliano (1995) developed a measure for a construct they call mood awareness. They conceptualize this variable as consisting of two major components – mood monitoring and mood labeling. Mood labeling is considered the healthier of the two and generally consists of noticing a particular mood and then labeling the mood without becoming overly concerned about the cause of the mood or its potential effect. Mood monitors, on the other hand, are very concerned with their mood, not only noticing their mood states readily, but also becoming preoccupied with their mood. For example, mood monitors might let the realization that they are in a bad mood particularly bother them and, as a result, focus much of their effort on attending to their mood rather than simply letting it go. Overall, Swinkels and Giuliano found that mood monitoring is associated with higher levels of negative affect (as measured on the PANAS; Watson, et al., 1988) and affective intensity (as measured on the Affect Intensity Scale; Larsen & Diener, 1987) while mood labeling is associated with positive affect (as measured by the PANAS) and higher levels of self-esteem (as measured by the Global Self-Esteem Scale; Rosenberg, 1965). Of the two, Swinkels and Giuliano (1995) state:

> Mood monitoring...should predict less desirable outcomes. Compared to mood labeling, mood monitoring may tax an individual's resources to a greater extent; that is, attention directed toward scrutinizing one's mood is attention taken from other concerns. Moreover, the tendency to scrutinize one's mood may induce negative affect...(p. 940).

In other words, mood monitoring (relative to mood labeling) is expected to consume cognitive resources. The consumption of these cognitive resources is expected to result in poorer athletic performance due to the attentional deficits induced by mood monitoring. Therefore, it is hypothesized that a player's tendency to engage in mood monitoring will be associated with poorer athletic performance and that the player's tendency to engage in mood labeling will be associated with better athletic performance.

Rumination is a broad class of thoughts that revolve around a common theme and recur in the absence of environmental demands requiring the thoughts (Martin & Tesser, 1996). Scott & McIntosh (1999) developed a trait measure (the *Scott-McIntosh Rumination Inventory*; SMRI) of the type of rumination classified as worry by Martin & Tesser (1996). Rumination, as measured by the SMRI, consists of three components; a distraction component, an affective component, and a motivation component. The distraction subscale of the SMRI indicates the degree to which an individual typically is distracted from other tasks by ruminative thoughts. The affective component of the SMRI indicates the extent to which an individual is typically made to feel emotionally labile by ruminative thoughts. Finally, the motivation subscale of the SMRI indicates the extent to which an individual typically finds that ruminative thoughts serve as a prompt to work harder at goals. Taken together, these subscales indicate the extent to which people experience a range of behavioral and emotional responses to thoughts regarding their goal pursuits. Scott and McIntosh (1999) found that people who tend to ruminate also experience more negative affect (as measured by the PANAS; Watson et al., 1988), more worry [as measured by both the *Penn State Worry Questionnaire* (Meyer, Miller, Metzger, & Borkovec, 1990) and the *Worry Domains Questionnaire* (Tallis, Eysenck, & Matthews, 1991)], and perform less well on cognitively demanding tasks. Due to the cognitive demands placed upon tennis players and the tendency for rumination to consume cognitive resources, it is hypothesized that rumination will be negatively correlated with the athletic performance of tennis players.

The sport of tennis is believed to be one of the more highly mentally demanding sports as its successful execution requires not only a high degree of concentration for proper execution of serves and ground strokes but also for planning the strategy of the game and attending to the movements of one's opponent (Kriese, 1993). At the same time, the successful tennis player must remain relatively calm and in control emotionally. Any stimulus that successfully competes for precious cognitive resources should negatively impact athletic performance. In the present research, three potentially cognitive resource-consuming variables were explored for their relationship to the athletic performance of tennis players. It was hypothesized that negative affectivity, rumination, and mood monitoring would be negatively correlated with athletic performance measures.

METHOD

Participants

Sixteen NCAA Division II intercollegiate tennis players participated in the season-long (February to May) study with no tangible reward. The players were 10 males and 6 females aged 19-26 ($M = 21.87$, $SD = 2.58$) years. Relative to their ages, the players, on average, had considerable competitive experience ($M = 12.33$ years, $SD = 3.8$ years). The academic classification of participants was as follows: 5 Freshmen, 3 Sophomores, 4 Juniors, and 4 Seniors. The sample consisted predominantly of international students (87.5%) most of whom were from European countries (71.4%).

Materials

Participants were given a battery of questionnaires at the beginning of the season including: the Scott-McIntosh Rumination Inventory (SMRI; Scott & McIntosh, 1999), the Mood Awareness Scale (MAS; Swinkels & Giuliano, 1995), and the Positive Affect Negative Affect Schedule (PANAS – general time-frame version; Watson et al., 1988). In addition, each participant completed a measure of post-match mood (now time-frame version of the PANAS; Watson et al., 1988) following the observation.

Procedure

Participants were asked to complete the initial battery of questionnaires during a team meeting at the beginning of the Spring season before the start of any intercollegiate competition. During this meeting, the general purpose of the research was described to the participants and informed consent was obtained. It was explained that during the regular season, players would be randomly selected for observation and that each player being observed would be asked to complete one post-match questionnaire. Participants were informed that they could withdraw their participation at any time without penalty. Following completion of the inventories at the initial meeting, participants placed their completed questionnaires in a manila envelope which was labeled with a uniquely assigned identification number and placed the envelope in a box with other completed packets that were then processed into aggregate files by an undergraduate research assistant who was unaware of each player's identity. Players were assured that their responses would be kept confidential.

Following the initial meeting, each player was randomly chosen for two observations of singles play over the course of the regular season. On the day of competition, each player was informed of his or her selection for observation at least 20 minutes before the beginning of the singles match (in order to allow the player to refuse participation if desired). At that time, the observer positioned himself or herself to record the match. The observer chose a position that would be about 10 feet from the court, viewing the court from the side with a point of view similar to that of a spectator (oftentimes sitting with the spectators). The observer was responsible for recording the point-by-point progress of the match using a system similar to that developed by Paul Scarpa (Kriese, 1993). This system involves writing a symbol to represent the stroke that ended the point with the addition of a subscripted number to indicate the order of the points in a particular game. In using this system, the entire match was recorded on paper, allowing easy analysis of the match for the first serve percentage, the number of unforced errors, winners and forced errors committed by both players during the match.

After the completion of the match, the player was approached by the observer and asked to complete the post match mood measure (i.e., the now version of the PANAS). The players completed these questionnaires while sitting alone recuperating from the match.

RESULTS

Athletic Performance Measures

For each athlete, 4 measures of athletic performance were developed from the observations of their singles match play in actual intercollegiate athletic competition during the Spring semester of 2000. Of the 16 players to complete the initial set of questionnaires, 10 were observed during play. Two of the six players were not observed because they refused to participate on the days they were selected. Two others were the equivalent of junior varsity players who did not compete during the season on a regular basis. The remaining two unobserved players quit the team before they could be observed. Of the 10 observed players, 8 were observed for four sets of tennis (two sets in each observed match). One player was observed for two matches, but during her second observation the match was halted following the first set due to the team having clinched the team victory. Therefore, only three sets of tennis were recorded for her. One player was only observed for one match (the equivalent of two sets). Because of the different number of sets of tennis observed for participants in the sample, the performance of each player was averaged across all observed sets, so that the athletic performance measure represents an average set for each observed player.

For each player, a measure of first serve percentage (TFSP) was determined by dividing the number of serves that were "good" (i.e., the technical term for the situation in which the ball landed in the proper service box without a let or fault being called) on the first attempt at service delivery for each point, divided by the total number of serves (i.e, the sum of first serves, second serves, and double-faults). A measure of forced error ratio (TFER) was calculated by summing the points won by forcing the opponent to make an error and dividing that number by the total number of points for the game (forced errors + winners + unforced errors). Finally, a ratio of winners to unforced errors (TWUER) was calculated by dividing the total number of winners by the total number of points that were not forced errors (winners + unforced errors). A composite measure of athletic performance (CPERF) was formed by summing the measure of first serve percentage (TFSP) and the ratio of winners to unforced errors (TWUER) and dividing by 2. CPERF (TFSP and TWUER) demonstrated high internal consistency with an alpha coefficient = .85.

A bivariate Spearman rank-order correlation matrix was produced by correlating the athletic performance measures with the proposed related variables (mood, rumination, and mood awareness).

Demographics

Sex of participant, number of years of playing tennis (EXP), and age were unrelated to any athletic performance measure, all *ps* > .27. Sex of participant was also unrelated to rumination, mood awareness, positive affect, or negative affect, all *ps* > .36. As a result, all remaining analyses are collapsed across the sex variable.

Reliability and Validity Evidence for Individual Difference Measures

The value of the present research necessarily depends upon the reliability and validity of the measures used. To examine the reliability of the scales used in the present study, internal consistency reliability coefficients for the individual difference measures were calculated and compared to the values observed in other samples. Comparisons were made to the internal consistency values reported in Swinkels and Giuliano (1995; Study 1) for Mood Awareness, Scott and McIntosh (1999; Study 1) for Rumination, and Watson et al. (1988) for Positive Affect (PAN) and Negative Affect (NAN). Values for the present study (with comparisons in parentheses) were as follows: Mood Monitoring (MM), .94 (.88); Mood Labeling (ML), .51 (.77); Rumination (RUM), .43 (.60); Emotion (EMOT), .78 (.68); Motivation (MOT), .73 (.77); Distraction (DIST), .25 (.68); Positive Affect, .63 (.88); and Negative Affect, .69 (.87).

Validity evidence was garnered by examining the construct validity of the various scales by exploring the theoretical relationships among the various scales. Previous work has indicated that Rumination, Mood Awareness, and Negative Affect are related constructs (Scott, 1999) and that Mood Awareness and Negative Affect are related (Swinkels & Giuliano, 1995). Therefore, observed correlations among these variables can be taken as validity coefficients indicating construct validity for the various scales. Spearman correlations were computed to provide a comparison of the relationships in the present study to those in previous samples. Mood Monitoring and Rumination were correlated, $r_s(14) = .56$, $p < .05$. Mood Monitoring and Negative Affect were correlated, $r_s(13) = .72$, $p < .002$. The emotionality subscale of the SMRI (EMOT) was correlated with Mood Monitoring, $r_s(14) = .60$, $p < .02$ and Negative Affect, $r_s(13) = .58$, $p < .01$.

As further evidence of the acceptability of the individual difference measures for the present research, the means and standard deviations of each of the measures were compared to the norm population parameters established with the development of each scale and reported in the literature. Comparisons were made using the SPSS one sample t-test analysis and substituting the population mean from the literature for the test value in the analysis. Only two variables were found to deviate significantly from the norm – Positive Affect and Negative Affect, both *ps* $< .01$. When population values from an athletic sample were substituted [taken from PANAS *M*s and *SD*s reported in Greenless, Nunn, Graydon, & Maynard (1999)] rather than the original norm values, only Negative Affect continued to be significantly deviant, $p < .005$.

Athletic Performance and Individual Difference Measures

Rumination. The emotionality subscale of the SMRI (EMOT) was negatively correlated with the winner to unforced error ratio (TWUER), r_s (8) = -.71, $p < .05$. Players who tend to be emotionally reactive to difficulty in reaching goals had a lower ratio of winners to unforced errors. The overall measure of rumination (RUM) was negatively correlated with the composite measure of athletic performance (CPERF), r_s (8) = -.63, $p < .05$. Players who report higher rates of trait rumination tend to have poorer athletic performance. None of the rumination measures related to first serve percentage (TFSP) or ratio of forced errors (TFER), all *ps* $> .28$.

Mood Awareness. Three of the four measures of athletic performance were found to positively correlate with the mood awareness subscale mood labeling (ML), the type of mood awareness that is generally associated with more healthy processing of mood; TFSP & ML, r_s (8) = .81, $p < .005$, TWUER & ML , r_s (8) = .79, $p < .01$, and CPERF & ML, r_s (8) = .81, $p < .005$, indicating that mood labeling is associated with better athletic performance. Mood monitoring (MM), the less desirable and generally less healthy form of mood awareness was found to be negatively correlated with TWUER, r_s (8) = -.65, $p < .05$. It appears that mood monitors tend to perform less well on the court.

Negative Affect. Negative affect measured as a global measure during the initial battery of questionnaires (NAN) was unrelated to any performance measure, all *ps* $> .10$. Negative affect measured at the first observation (ONAN) was found to correlate negatively with TFSP, r_s (8) = -.67, $p < .05$ and CPERF, r_s (8) = - .72, $p < .05$. Players who did not play well reported more negative affect in post match measures.

Positive Affect. Preseason positive affect (PAN) was unrelated to any performance measure, all *ps* $> .69$. Positive affect measured at the first

(OPAN) and second observations (SPAN) was unrelated to any performance measure, all *ps* > .09.

DISCUSSION

Athletic performance was related to rumination, mood, and mood awareness. Mood awareness clearly demonstrated the strongest relationship in the present study (particularly the subtype, mood labeling). It appears that being a mood labeler may be of substantial benefit for athletic performance. Of course, further investigation of the observed relationship between mood labeling and athletic performance is necessary before any definitive conclusions can be drawn regarding the nature of the relationship. It is, however, a very promising avenue that warrants more research.

This study has its share of limitations, and we would be remiss if we did not address those problems directly. First and foremost is the limited sample and sample size. Admittedly, the sample is smaller than we would have liked for this study. With a very limited initial sample size, it did not take long for the sample to dwindle to a relatively low 10 useable participants. To address this problem, we utilized the Spearman rank-order coefficient to help minimize the impact of the limited sample size (and any resulting departure from normality). Additionally, because our sample is a group of highly trained (perhaps elite) athletes, we have taken pains to limit the tendency to try to generalize the results of this research beyond the scope of the sample we are working with. We recognize that the external validity of the research may be questioned if one were to try to apply the findings to a markedly different group of players (e.g., recreational players), so we do not advocate such broad generalizations. In our own follow-up research, we have retained the same sample of players (i.e., a sample of elite intercollegiate tennis players) with the intention to attempt to broaden the external validity of the study to other groups (recreational players, athletes in other sports) at a later date.

A second limitation of the study is that there was no record kept of a given player's doubles performance and the potential for carryover effects of that performance from doubles to singles competition. It is possible that on any given team competition day, players may have allowed their performance in doubles to affect their performance in singles. In order for carryover effects to be a problem in the analysis of the data, there would have to be a systematic bias present – a tendency for one or more players to be overly concerned about their failures (or successes) in doubles. Otherwise the variable can be considered a random variable and should, if anything, generate only random "noise" in the data. An investigation of writing samples taken following singles competition does not indicate any evidence of any player focusing on doubles competition during their singles play. Thus, in the current research, the performance of any given player in doubles competition was considered a random, extraneous variable.

A final limitation of the study is that the PANAS scores for both positive and negative affect indicate departure from the norms obtained in the norming samples, thus bringing into question the utility of this measure for athletes. Even when compared to a specific athletic sample, mean negative affect in this sample continued to show significant departure from normality. These results are not surprising given the difficulty of previous researchers to find utility in mood measures to differentiate athletic performance. The concerns over this issue for the present research are minimal given that the purpose of the present research was to explore alternatives to the mood—athletic performance relationship. In that vein, it is important to note that the means on the other scales are consistent with the expected norms from the norming samples.

While a great deal of research and therapeutic effort have been exerted in an attempt to better understand and alleviate the deleterious effects of negative mood states and related variables (such as anxiety) in athletes (e.g., Anshel & Wells, 2000; Rawstorne, Anshel & Caputi, 2000; Anshel, Jamieson, & Raviv, 2001), many questions have been left unanswered. We believe that the field of social psychology (particularly in the area of understanding mood and its effects on behavior) can make a meaningful contribution in addressing these questions. For example, Swinkels and Guilliano (1995) have proposed that one potential use of the MAS (and its ability to differentiate people as mood labelers versus mood monitors) would be to apply the inventory in a clinical setting. In that setting it could be used to help both therapists and clients discern the ability of the client to facilitate mood regulation. They imply that mood labeling might be "taught" as a regular part of therapy. By extension, the present research seems to point to the same conclusion for tennis players. Without being overly optimistic, the MAS and perhaps to a lesser extent, the SMRI, might be employed by sport psychologists and coaches alike as tools to aid them in helping and assessing their athletes. Once again, recognizing the importance of further study in this area, we advocate that other researchers employ these tools in their own investigations of the variables related to athletic success.

REFERENCES

Anshel, M. H., & Wells, B. (2000). Personal and situational variables that describe coping with acute stress in competitive sport. *Journal of Social Psychology, 140*, 434-450.

Anshel, M. H., Jamieson, J., & Raviv, S. (2001). Cognitive appraisals and coping strategies following stress among skilled competitive male and female athletes. *Journal of Sport Behavior, 24*, 128-143.

Beedie, C. J., Terry, P. C., & Lane, A. M. (2000). The profile of mood states and athletic performance: Two meta-analyses. *Journal of Applied Sport Psychology, 12*, 49-68.

Friend, J., & LeUnes, A. D. (1990). Predicting baseball player performance. *Journal of Sport Behavior, 13*, 73-86.

Gould, D., & Udry, E. (1994). Psychological skills for enhancing performance : Arousal regulation strategies. *Medicine and Science in Sports and Exerc*ise, *26*, 478-485.

Greenless, I. A., Nunn, R. L., Graydon, J. K., & Maynard, I. W. (1999). The relationship between collective efficacy and precompetitive affect in rugby players: Testing Bandura's model of collective efficacy. *Perceptual & Motor Skills, 89*, 431-440.

Kriese, C. (1993). *Winning Tennis*. Indianapolis, IN: Master's Press.

Larsen, R. J., & Diener, E. (1987). Affect intensity as an individual differences characteristic: A review. *Journal of Research in Personality, 21*, 1-39.

Martin, L. L., & Tesser, A. (1996). Some ruminative thoughts. In R. S. Wyer, Jr. (Ed.), *Advances in social cognition*. (pp. 1-47). Hillsdale, N J: Lawrence Erlbaum Associates.

Meyer, T. J., Miller, M. L., Metzger, R. L., & Borkovec, T. D. (1990). Development and validation of the Penn State Worry Questionnaire. *Behaviour Research and Therapy, 28*, 487-495.

Newby, R. W., & Simpson, S. (1994). Basketball performance as a function of scores on Profile of Mood States. *Perceptual and Motor Skills, 78*, 1142.

Prapavessis, H. (2000). The POMS and sports performance: A review. *Journal of Applied Sport Psychology, 12*, 34-48.

Rawstorne, P., Anshel, M. H., & Caputi, P. (2000). Exploratory evidence of coping styles following acute stress in sport. *Australian Journal of Psychology, 52*, 75-82.

Renger, R. (1993). A review of the Profile of Mood States (POMS) in the prediction of athletic success. *Journal of Applied Sport Psychology, 5*, 78-84.

Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.

Rowley, A. J., Landers, D. M., Kyllo, L. B., Etnier, J. L. (1995). Does the iceberg profile discriminate between successful and less successful athletes? A meta-analysis. *Journal of Sport and Exercise Psychology, 17*, 185-199.

Scott, V. B., Jr., & McIntosh, W. D. (1999). The development of a trait measure of ruminative thought. *Personality and Individual Differences, 26*, 1045-1056.

Scott, V. B., Jr. (1999, March). *Rotten moods and brighter days: The benefits of mood labeling*. Poster session presented at the annual meeting of the Southeastern Psychological Association, Savannah. GA.

Swinkels, A., & Giuliano, T. A. (1995). The measurement and conceptualization of mood awareness: Monitoring and labeling one's mood states. *Personality and Social Psychology Bulletin, 21*, 934-949.

Tallis, F., Eysenck, M., & Mathews, A. (1991). A questionnaire for the measurement of nonpathological worry. *Personality and Individual Differences, 12,* 21-27.

Terry, P., Coakley, L., & Karageorghis, C. (1995). Effects of intervention upon precompetition state anxiety in elite junior tennis players: The relevance of the matching hypothesis. *Perceptual and Motor Skills, 81,* 287-296.

Treasure, D. C., Monson, J., & Lox, C. L. (1996). Relationship between self-efficacy, wrestling performance, and affect prior to competition. *The Sport Psychologist, 10,* 73-83.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54,* 1063-1070.

Authors' Notes We would like to thank Simon Earnshaw, the AASU tennis coach, and the AASU tennis teams for their support in making this research possible through their participation. We would also like to thank John Kraft and Jonathan Roberts for their comments on an earlier draft of this article and Geraldine McLean for her assistance in preparation of the manuscript. Portions of this research were presented at the Annual Meeting of the Southeastern Psychological Association on March 23rd, 2002, in Orlando, FL.

Factors Affecting Vote Decisions: An Example from the Alabama Lottery

Larry Powell
and
William R. Self
University of Alabama at Birmingham

In 1999, the state of Alabama held a statewide referendum to decide whether to legalize a lottery system as a means of raising state funds. The campaign featured two groups -- proponents who were led by government and business leaders and opponents who organized around a coalition of religious groups. This study analyzed the complexity of the political attitudes of voters in that referendum, comparing the factor structure of pro-lottery voters, anti-lottery voters, and undecided voters. The analysis indicated that pro-lottery voters viewed the referendum from a single-factor structure dominated by financial considerations. Anti-lottery voters viewed the issue from a two-factor perspective - financial considerations versus moral considerations. Undecided voters had the most complex factor structure of the three; in addition to financial and moral considerations, they also viewed the issue as having a social dimension. The results are discussed in terms of campaign strategy and implications for understanding undecided voters.

Religious attitudes can have an impact on any number of political issues. Research on attitude formation has suggested that attitudes can often be a reflection of an individual's personal values, including religious values. Katz (1960), for example, identified a value-expression function of attitudes in which "the individual derives satisfactions from expressing attitudes appropriate to his personal values and to his concept of himself" (p. 170). In essence, value-expressive attitudes reflect individuals' self-perceptions of who they are and their roles within an external world, allowing them to give "positive expression to his central values and to the type of person he conceives himself to be" (p. 173).

One area in which values can have a major impact on attitudes is within the political arena. Jalen (1988) reviewed two decades of scholarship on religion and mass political behavior and demonstrated that religious affiliations and beliefs affect a wide variety of political attitudes and activities. Leege and Kellstedt (1993) found that religion has influenced the

---

Author info: Correspondence should be sent to: Dr. Larry Powell, University of Alabama at Birmingham (205-934-3877), 901 South 15th Street, Birmingham, Alabama 35294-2060.

North American Journal of Psychology, 2002, Vol. 4, No. 3, 469-478.

© NAJP

political behavior of large groups of voters. Such findings are not surprising, considering that political ideologies are inherently value-based (liberal versus conservative) and can influence a number of different issues. Some of the nation's religious groups have recognized this factor and have invoked television and radio to express their attitudes and political goals. Abelman (1994) contended that "The Religious Right turned the electronic pulpit into a forum for public morality and political renewal" (p. 888). Podesta and Kurtzke (1990) asserted that the Religious Right has used mass communication when attempting to influence decision-makers to make legal and education systems reflect public morality. In the Alabama Lottery referendum, both sides used the electronic media extensively.

Other observers have noted that religious values can also play a vital role in political attitudes. Lopatto (1985) argued that theology, based on religious affiliation, served as a sieve through which voters filtered and evaluated their political worlds. That phenomenon was a major factor during the 1980s with the participation of the Moral Majority (Wald,1987) and continued during the 1990s (Manza & Brooks, 1997) through groups like the Christian Coalition. Religious values can potentially have a double impact on attitudes, since those values are often related to both self-perception and group identification (with a church or religious organization). Group identification is a known factor in framing perceptions (Penning, 198[illegible]). Katz (1960) argued that identification with a religious organization could influence people to internalize the values of the organization, and the presence of a cue associated with those values can motivate them to vote for or against a particular candidate (p. 187). That concept was supported by Wald (1987), who found that integration into religious groups influences both the political attitudes and political behavior of the groups' members. In previous U.S. elections, for example, religious groups have effectively used "family values" as a decisional motivator for church members (Hammond, Shibley, & Solow, 1994).

But religious values, even among the religious, are not the only factors influencing attitude formation. Another factor that influences political attitudes is candidate choice. Once a voter has decided to support a particular candidate, that decision serves as an attitudinal anchor that is used to evaluate other messages relevant to the campaign. The campaign process tends to polarize perceptions (Nimmo & Savage, 1976), increasing perceptions that the opposing candidates have different images (Strickler, 1964), and leading to harsher views of the candidates that the voter opposes (Raven & Gallo, 1965). Further, vote decision can have a direct effect on the structure of such attitudes. Powell (1977), for example, found that voters had a more distinct, more differentiated, and more complex image of the candidate they were voting against than of the candidate they supported.

Political anchors are not limited to candidates. Religious group members may also be activated by an issue related to some particular religious value (Kniss, 1996). While many political issues might have the capacity to serve as a value cue, some, such as gambling, seem to activate more religious values than others. Many religious groups oppose any type of gambling or games of chance, and ballot issues related to gambling can activate attitudes based on religious values. Those attitudes can have a major impact on voting behavior because they so directly reflect the "most fundamental and cherished assumptions" about how people will live their lives (Hunter, 1991, p. 42). The lottery seems to be such an issue; Ellison and Nybroten (1999), for example, found that anti-lottery views were more prevalent among Baptists and other conservative religious denominations than among the public as a whole.

This study examined the role that religion had in attitude formations related to a lottery referendum in Alabama. The pro-lottery position argued that the lottery would provide a financial base for improving education in the state. They were opposed by a coalition of religious groups that argued the lottery was a form of gambling and therefore morally wrong (Ellison & Nybroten, 1999; Gravatt, 1999). As the campaign progressed, the following issues and arguments emerged: concerns over continuing lost lottery revenue to neighboring lottery states, questions about the source of funding of the anti-lottery side, and questions about the impact of the lottery on the lifestyles of residents in the state (Gravatt, 1999). At the outset of the campaign - six months prior to the election - public opinion was clearly on the pro-lottery side; early polls in January before the election indicated that nearly 70 percent of the voters intended to support the referendum (Parker Group, 1999). As the election neared, support for the lottery had decreased to a thin majority (Gravatt, 1999), and the issue was ultimately defeated by 54 percent-to-46 percent margin.

The purposes of this study were to (1) verify whether people who frequently attended church were more likely to oppose the lottery and 2) identify whether the attitudinal structure of those voters who were undecided late in the campaign was more complex than those voters who were already committed to either the pro-lottery or anti-lottery side. Results from previous research would predict that pro-lottery voters would have a more simplistic attitudinal structure than those who opposed the lottery (Powell, 1977), but little empirical information is available to make predictions about the attitudinal structure of undecided voters. Would the attitudinal structures of undecided voters on this issue be more or less complex than those of decided voters? Arguments can be made for both sides. DeVries and Tarrance (1972), for example, argued that undecided voters have a more complex understanding of issues and candidates; that is, they are undecided because they recognize the complexity of the issues and have difficulty making a decision. Conversely, Shadegg (1972) contended that many undecided voters have a simplistic notion of candidates and issues. They are undecided because they lack the information or motivation to make a decision.

The authors hypothesized that frequent churchgoers would be more likely to vote against the lottery, while non-churchgoers would support the lottery.

What are the differences in the factor structure of attitudes toward a statewide lottery of voters who support a lottery, those who oppose the lottery, and those who are undecided? The authors hypothesized, based on the previous research by Powell (1977), that the factor structure of pro-lottery voters would be less complex than those of anti-lottery voters.

If DeVries and Tarrance (1972) were correct, the attitudinal structure of undecided voters should be more complex than those of either the pro-lottery or anti-lottery voters. If Shadegg (1972) was correct, the attitudinal structure of undecided voters should be less complex than either of the others.

METHOD

Participants

Participants in the study were 400 registered voters in Alabama who were interviewed in a statewide random sample taken September 27-30, 1999. Participants were selected using the following procedure. First, a sampling formula was developed by dividing the state into six geographical areas based upon the state's six television media markets. Stratification for each area was based upon voter participation in the 1996 presidential election. Second, 8,000 names matched with phone numbers were randomly selected and divided into subgroups based on the geographical breakdown. Names were randomly selected, and individuals were contacted by telephone and asked if they were registered voters and if they would participate in the research project. The response rate for those who qualified was 71.4%.

Instrument

The survey questionnaire consisted of 14 questions regarding the October 12, 1999 Alabama lottery referendum. Voting intention was measured with the following question: "On October 12, a referendum will be held on establishing a statewide lottery in Alabama. If that referendum were held today, would you vote Yes, for the lottery, or No, against the lottery?" "Undecided" was coded as the third option.

Attitudes toward lottery issues were measured with a series of six five-interval Likert-type questions (responses ranging from "strongly agree" to "strongly disagree"). Each statement was drawn from arguments used by the campaign organizations, in their brochures, campaign ads, or press releases. The specific questions asked were:

1. "The lottery is needed to help finance education." (*Finance education*)
2. "The lottery is needed in Alabama so that gambling revenues will not be lost to other states such as Georgia and Florida." (*Avoid Loss of Revenues*)
3. "Citizens of Alabama should have a chance to play the lottery and become millionaires just like people in Georgia and Florida." (*Millionaires*)
4. "Gambling is morally wrong." (*Morally wrong*)
5. "Passing a lottery in Alabama will lead to other types of gambling." (*Lead to more gambling*)
6. "Poor people bear the financial burden of the lottery." (*Poor bear burden*).

Frequency of church attendance was measured with the following question: "Do you usually attend church once a week or more, a few times a month, or a few times a year or less?" Those who said "once a week or more" were defined as frequent church attenders. Those who reported attending "a few times a month" were defined as occasional church attenders. Those who answered "a few times a year or less" were defined as non-church attenders.

The remaining six questions included were demographic questions (age, income, party orientation, race, gender, and geography) used to ensure that the sample was not biased toward any specific demographic group.

The participants in the survey were divided into three subgroups based on their expressed voting intention: pro-Lottery voters ($n = 192$), anti-lottery voters ($n = 120$), and undecided voters ($n = 88$). The attitudinal responses of the participants in each individual group were analyzed with a factor analysis using a principal component analysis.

Statistical Analysis

Two statistical analyses were conducted on the resulting data. The first research question was tested with a Chi-Square analysis. The second research question was tested with a principal component factor analysis with a varimax rotation to identify the factor structure; a minimum eigenvalue of 1.00 was used as the criteria for establishing the presence of a factor in both analyses.

RESULTS

Impact of Religion on Vote Choice

As predicted, frequent churchgoers were more likely to be anti-lottery voters, while non-churchgoers were more likely to be pro-lottery voters ($X^2_{1,6} = 72.08, p < .001$). Most (60.7%) of those voters who attended church weekly were anti-lottery voters. Most of those who rarely attended (74.4%) or who only occasionally went to church (63.0%) were pro-lottery voters.

Factor Structures.

The factor analysis on the responses of the pro-lottery voters identified a one-factor solution that accounted for 36% of the variance. The principle component matrix revealed the coefficient structure set out in Table 1.

TABLE 1 Factor Structure: Pro-Lottery Voters on Left, Anti-Lottery Voters in Center, Undecided Voters on the Right.

	Factor	Factors		Factors		
Items	1	1	2	1	2	3
Finance education	*.64*	*.70*	*.40*	*.74*	*.40*	*.22*
Avoid loss of revenue	*.69*	*.71*	*.55*	*.67*	*-.16*	*-.12*
Millionaires	*.66*	*.67*	*.39*	*.69*	*.25*	*.01*
Morally wrong	.55	*.40*	*-.68*	*.29*	*.81*	*.30*
Lead to more gambling	-.57	*-.71*	*.38*	*-.32*	*.78*	*-.05*
Poor bear burden	-.47	*-.51*	*.52*	*-.40*	*.04*	*.83*
eigenvalues	2.16	2.51	1.61	1.99	1.32	1.01

The factor analysis on the responses of the anti-lottery voters identified a two-factor solution that accounted for 69% of the variance. The principle component matrix revealed the factor coefficient structure shown in the center of Table 1.

The factor analysis on the responses of the undecided voters identified a three-factor solution that accounted for 72% of the variance. A principle component matrix revealed the factor coefficient structure shown on the right in Table 1.

DISCUSSION

Three significant results were identified. First, opposition to the lottery was associated with church attendance. A majority of those voters who attended church weekly were anti-lottery voters, while most who rarely or only occasionally attended church were generally pro-lottery voters. This result supports the previous research by Ellison and Nybroten (1999) and indicates that initial support and opposition to the lottery were heavily influenced by religious orientation.

Second, as measured by the above questionnaire, anti-lottery voters had a more complex attitudinal structure related to the lottery than did pro-lottery voters. Pro-lottery voters saw the referendum as a one-dimensional concept that focused on the financial aspects of the issue (i.e., the lottery would provide large sums to the state coffers). The pro-lottery position was attractive to these voters because it could avoid future taxes. Conversely, anti-lottery voters viewed the referendum as a two-dimensional issue - the positive financial aspects of the lottery versus the negative moral values associated with gambling. This aspect of the study supports Powell's (1977) earlier finding that anti-voting attitudinal structures are more complex than pro-voting attitudes.

Third, undecided voters had the most complex attitudinal structure - a three-dimensional construct that included the social ramifications in addition to moral and economic considerations. There is research that indicated that some church attendees would be in this group, as many churchgoers are reluctant to allow their religious beliefs to determine their political decisions (Wolfe, 2000). If so, though, these voters still viewed the lottery referendum as a more complex issue than did other churchgoers.

This finding has several implications related to general communication principles and to the specifics of this particular campaign. On a general level, the implication that undecided voters have complex political attitudes is consistent with DeVries and Tarrance's (1972) assumptions about swing voters. This is somewhat surprising, considering that subsequent studies on split-ticket voting have raised questions about the validity of DeVries and Tarrance's original descriptions of these voters. Subsequent research has generally indicated that swing voting is a byproduct of campaign targeting rather than a personality characteristic of particular voters (Powell & Kitchens, 1994). If this is true, then the complexity of the undecided may be a byproduct of such target communication. Targeted communication may contribute to the complexity of attitudes among undecided voters. Each side targets undecided voters with persuasive messages; their processing of that information creates a more complex understanding of the issue than either the pro- or anti-sides would have by themselves. That process doesn't totally explain the complexity of undecided voters, though, because the undecided voters added a social dimension to their structure that did not exist independently for either the pro- or anti- sides. The added complexity of attitudes among undecided voters may be the factor that makes them remain undecided for a longer period during the campaign.

On a campaign-specific level, many local observers attributed the defeat of the lottery to a strong stand by the religious community. However, the analysis of these data indicates that the anti-lottery side would have been unsuccessful if its position had been limited to association with religious values. Although there was a relationship between church attendance and anti-lottery voting, that relationship was not strong enough to have defeated the lottery without additional anti-lottery votes. Only about 60 percent of the frequent church-goers were against the lottery; while that represents a majority, it is far from the overwhelming numbers that might have been anticipated.

Religious activity may be conceptualized in various ways. For example, extrinsic religiosity features regular church attendance, following religious prohibitions, and observing religious holidays, among other outward manifestations. Intrinsic religiosity, on the other hand, concerns the carrying out and adherence to a religious belief system (Bergin, 1983). This study measured extrinsic religiosity. As such, conclusions about the role of intrinsic religiosity cannot be directly drawn from these data. Still, the modest anti-lottery vote among church attendees may reflect the declining significance of denominalizationism (Wuthnow, 1988; Hunter, 1991). This approach asserts that social differentiation within religious denominations, rising levels of religious mobility and intermarriage, and declining organizational conflicts have decreased the political relevance of denominational membership. Such social differentiation may be a byproduct of increased secularization of religion. The "secularization thesis" argues that the growing secularization of citizens is a source of declining political influence of religion (Hout & Greeley, 1987; Inglehardt, 1999). If so, the inverse could also be occurring, with religion having a declining influence on politics. Such a possibility has led one observer to conclude that: "no matter how deep their faith commitments, Americans are reluctant to allow their beliefs to determine their political outlook" (What Scholarship Reveals, 2000, p. B7).

Regardless, those who voted against the lottery for religious reasons were a significant group, just not enough to win the election. The swing voters who decided the election were those undecided voters who ultimately voted against the issue. Their attitudinal structure was more similar to anti-lottery respondents than to pro-lottery respondents, and most of them ultimately moved in that direction. But the presence of a third factor related to social ramifications implies that the arguments of the anti-lottery side had to expand beyond religious arguments to be successful.

Still, these results should be interpreted with caution, since several limitations may inhibit the generalizations of the results. One potential limitation has to do with geography. This election was held in Alabama, a southern state in the middle of the South's "Bible Belt." As such, the pro-religious and anti-lottery attitudes of this area may be stronger and more complex than those in other areas. After all, most other states - including nearby Georgia - have successfully passed lottery referenda even in the face of opposition from religious groups. Second, this was a special referendum with no other major element on the ballot. There were no opposing candidates for any elective office, resulting in the lottery being the sole focus of all media discussion and advertisement. This solitary focus could have intensified the attitudinal complexity of this issue, particularly for undecided voters.

In future dichotomous political battles, it would be useful to investigate whether the religious side recognized the attitudinal complexities of the issues, and, if so, whether it structured any of its messages to influence undecided voters by mentioning the social impact factor. Such ideas raise further questions about the complexity of attitudes among undecided voters. Future study should seek to identify whether the attitudinal structures of undecided voters in candidate-based campaigns display the same degree of complexity as found in this issue-specific campaign.

REFERENCES

Abelman, R. (1994). News on the 700 Club: The cycle of religious activism. *Journalism Quarterly, 74,* 887-892.

Bergin, A. E. (1983). Religiosity and mental health: A critical reevaluation and meta-analysis. *Professional Psychology: Research and Practice, 14,* 170-184.

DeVries, W., & Tarrance, V. L. (1972). *The ticket-splitter.* Grand Rapids, MI: William B. Eerdmans.

Ellison, C. G., & Nybroten, K. A. (1999). Conservative Protestantism and opposition to state-sponsored lotteries: Evidence from the 1997 Texas poll. *Social Science Quarterly, 80(2),* 356-369.

Gravatt, R. (1999, October 5). Two poll results at odds: UAB survey says most approve. *Birmingham Post-Herald*, pp. E1-E2

Hammond, P. E., Shibley, M. A., & Solow, P. M. (1994). Religion and family values in presidential voting. *Sociology of Religion, 55,* 277-290.

Hout, M., & Greeley, A. (1987). The center does not hold. *American Sociological Review, 52,* 325-345.

Hunter, J. D. (1991). *Culture wars.* New York: Basic Books.

Inglehart, R. (1990). *Culture shift.* Princeton, NJ: Princeton University Press.

Jalen, T. G. (1988). Research in religion and mass political behavior in the United States: Looking both ways after two decades of scholarship. *American Politics Quarterly, 26,* 110-134.

Katz, D. (1960). The functional approach to the study of attitudes. *Public Opinion Quarterly, 24,* 163-204.

Kniss, F. (1996). Ideas and symbols as resources in intrareligious conflict: The case of American Mennonites. *Sociology of Religion, 57.* 7-23.

Leege, D. C., & Kellstedt, L. A. (1993). *Rediscovering the religious factor in American politics.* Armonk, NY: M.E. Sharpe.

Lopatto, P. (1985). *Religion and the presidential election.* New York: Praeger.

Manza, J., & Brooks, C. (1997). The religious factor in U.S. presidential elections, 1960 B 1992. *American Journal of Sociology, 103,* 38-81.

Nimmo, D., & Savage, R. L (1976). *Candidates and their images.* Pacific Palisades, CA: Goodyear.

The Parker Group (1999). Alabamians call for quick implementation of a state lottery. *Alabama Insight Poll*, 1.

Penning, J.M. (1988). The political behavior of American Catholics: An assessment of the impact of group integration vs. group identification. *Western Political Quarterly, 41 (2)*, 289-308.

Podesta, A. T., & Kurtzke, J. S. (1990). Conflict between the electronic church and state: The religious right's war against pluralism. In R. Abelman & S. Hoover (Eds.), *Religious television: Controversies and conclusions* (pp. 207-226). New York: Ablex Publishing.

Powell, L. (1977). Voting intention and the complexity of political images: A pilot study. *Psychological Reports, 40*, 243-246.

Powell, L., & Kitchens, J. T. (1994, June). Ticket splitting: Dead or alive. *Campaign & Elections*, pp. 34-35

Raven, B. H., & Gallo, P. S. (1965). The effects of nominating conventions, elections, and reference group identification upon the perception of political candidates. *Human Relations, 18*, 217-229.

Shadegg, S. (1972). *The new how to win an election*. New York: Taplinger.

Stricker, G. (1964). The operation of cognitive dissonance on pre- and post-election attitudes. *Journal of Social Psychology, 63*, 111-119.

Wald, K. D. (1987). *Religion and politics in the United States*. New York: St. Martin's Press

What scholarship reveals about politics and religion. (2000). *Chronicle of Higher Education, 47*, B7-B10.

Wolfe, A. (2000, Sept. 8). What scholarship reveals about politics and religion. *Chronicle of Higher Education, 47*, B7-9.

Wuthnow, R. (1988). *The restructuring of American religion: Society and faith since World War II*. Princeton, NJ: Princeton University Press.

An Interview with Sonia Nieto About Multiculturalism

(Interviewed on behalf of NAJP by)
Bill Gaedke
Michael F. Shaughnessy
Eastern New Mexico University

NAJP: *Affirming Diversity* is now in its third edition. How have things changed since the first edition, and how has the book been received?

SN: I started writing *Affirming Diversity* about 12 years ago, and the first edition was published in 1992. I've been overwhelmed by the positive response the book has received over the years. For instance, it has been very favorably reviewed in many journals that usually review only "scholarly " books rather than textbooks, which are viewed as having less status. At the same time, although it's mostly been used as a textbook in colleges and universities around the country, it's also been used in professional development courses in school districts, and in such unlikely places as medical schools, publishing companies, and training institutes. I've been gratified by the many, many letters, phone calls and in the recent past, e-mail messages that I've received from people around the world who've communicated with me about the book. Not all the messages have been positive of course, although a surprisingly high percentage of them have been. But, I've enjoyed all the comments I've received because they've taught me to think more critically about what I write. Many of the suggestions readers have made have found their way into the second and third and fourth editions through the changes I've made and if I'm fortunate enough to have the book in future editions, I'm sure that readers' comments will find their way into those editions as well.

NAJP: It is the year 2002 and we are nearing 2003! Where has multicultural education been in the last 20-30 years, and where do you see it going in the next 20 to 30 years?

SN: Multicultural education has grown tremendously since its beginnings about 30 years ago. I still remember walking into the first course I ever took in multicultural education. It was a graduate course taught by

---

Author info: Correspondence should be sent to: Dr. Michael Shaughnessy, Eastern New Mexico University, Portales, NM 88130.

North American Journal of Psychology, 2002, Vol. 4, No. 3, 479-488.
© NAJP

Professor Bob Suzuki, who became one of my mentors, and it was the first course I took as a newly admitted doctoral student. Taking that course changed my life. Now mind you, many of the ideas we talked about in the course were not totally new to me. After all, my own background as a Puerto Rican child born and raised in New York City had taught me a great deal about differences, and becoming a bilingual teacher had augmented what I knew from my personal childhood experience. But in multicultural education, I found a language to describe what I had been experiencing and I found people who were thinking and writing about these issues, people who thought these ideas were important. It was exhilarating for me!

Since its beginnings in the early 1970's, one of the major changes I've seen in the field is that it has become more inclusive. What began as an outgrowth of the civil rights movement, and especially as a call for equal educational opportunity for African American students, has expanded a great deal since those days. Those who fought for multicultural education recognized that African American children were being systematically deprived of a high quality education based on the inequality they experienced because of their race, and, that without the education they deserved, the "American Dream" was out of reach for most youngsters. We need to keep those roots in mind; they give the field its heart and soul. At the same time, we also need to recognize that many other young people are poorly served by the schools. This includes females and males, as well as students of language minority backgrounds and different races and ethnicities, students from various social groups and economic classes, and those with different sexual orientations and abilities. It also includes students of European background, who are likewise miseducated if they receive an education that doesn't prepare them to function in the twenty-first century, with its tremendous diversity. Generally speaking, multicultural education now includes them all. We can see this change, for example in how NAME * has grown since it was founded about twelve years ago, not only in size, but also in inclusion of different perspectives, different points of view and different issues. Also, when multicultural education first came on the scene there was an emphasis on curriculum integration, and this was understandable because the curriculum, the textbooks, all the classroom materials, and even children's books were overwhelmingly mono-cultural. I believe that now there's a greater awareness that, in spite of the importance of curriculum, there's more to multicultural education than simply changing the curriculum. More people are looking at multicultural education as a trans-formative project that will change the schools, not only in curriculum but also in climate, values and expectations.

Where do I see multicultural education going in the next 20-30 years? That's hard to know, of course. But, I'd venture a guess about at least one change, and it has to do with teacher education. The demographics in our country are not likely to ever revert to what they were. The United States is more a diverse nation than ever, and in the next fifty years or so, Whites will probably become the minority. Given this situation, it is clear to me that schools need to change how they do things, and because teachers are the engines that keep schools going, they too need to change. This situation has very particular consequences for teacher education. For instance, it means that all future teachers, not just bilingual or ESL or special education teachers, need to know more about students of different backgrounds and how to teach them most effectively.

It also means that teachers need to be aware of the ways in which our schools, rather than focusing on the potential that all children have for learning, have consistently failed some students because of their social identities.

Related to this, future teachers need to understand that it's no accident that the structural inequality in our society also finds its way into schools. We see this every day in terms of the students who have access to gifted and talented programs compared to those who end up in special education, or students who are encouraged to take advanced courses compared to those who are suspended and expelled at disproportionately higher rates.

I don't think we've done a very good job of preparing prospective teachers to work with students of diverse backgrounds, or to understand the racism and inequality that have been embedded in our schools since they first opened their doors. I hope that the growing attention to multicultural teacher education will be one major change in the years ahead.

NAJP: Who has influenced you and how?

SN: Many people have influenced me over the years, and many others continue to help me grow. As a young teacher, I learned from the many other teachers with whom I worked at P.S. 25, the first bilingual school in the Northeast. Hernán La Fontaine, the first principal of the school and one of the very first Puerto Rican principals in the city, was a role model for me. So was Antonia Pantoja, a legend in the Puerto Rican community, who, beginning in the 1960's demanded better education for Puerto Rican students and she continued to do so until her death in May 2002, at the age of 80. She will be missed!

I left the bilingual school when I was recruited as a young faculty member in the Puerto Rican Studies Department at Brooklyn College. It was there that I met and formed a life-long friendship with María Sánchez, the person who became the city-wide symbol of student and faculty self-determination because the administration at the College refused to name her the chair of the department in spite of the great support she had from students, faculty, and community leaders. After many building take-overs and demonstrations and arrests, we won our demands and María served as chairperson with distinction for almost twenty years. Those were heady days and I learned many important lessons about what is worth fighting for. María Sánchez died in December of 2000, leaving a wonderful legacy behind.

Many mentors at the University of Massachusetts were instrumental in helping me form my initial ideas about equity, diversity, and multicultural education. My dear friends and colleagues Dora and Luis Fuentes, who later moved to Massachusetts, were an inspiration to me in the 1970's during the struggle for equal education in New York City. There was Paulo Freire, of course; I fell in love with this man when I first read his books during my doctoral program in the mid-1970's. Later, as a new assistant professor, I was fortunate to work with Paulo in person for a number of years when he was a visiting scholar every January at the University of Massachusetts.

When I read his books, or heard him speak, it was just like lightning. Ideas that I had thought about but hadn't been able to articulate began to make sense: that education is always political, that the decisions we make as teachers can have consequences for our classrooms and for the students we teach, that education serves either to liberate or to domesticate students. These were exciting ideas, and they had a profound effect on me. Many wonderful colleagues over the years have also had an impact on my thinking, including all the scholars who work has guided and illuminated my own thinking. They and the thousands of students I've been privileged to teach over the past 20 years, are simply too numerous to mention here. I continue to learn a lot from teachers, especially teachers who are strong and passionate and who struggle every day to do the best they can for their students. This is in fact, the subject of my latest book, *What Keeps Teachers Going in Spite of Everything?* (Teachers College Press, in press).

One of the best teachers I've ever had is my husband, Angel, who has taught me so many things about good teaching, and about what it means to love your students. And, probably, more than anything else, becoming a mother, and later a grandmother, makes the issues I think and write about much more consequential for me in an intensely tangible way.

NAJP: What training do teachers need to deal with the issues you raise in your books?

SN: In *Affirming Diversity: The Sociopolitical Context of Multicultural Education* I tackled what multicultural education means. I developed twelve case studies of young people of various backgrounds—from interviews by some colleagues and myself—to explore how the young people experienced schooling and how their backgrounds, their families, and the schools they attended might influence their education. In my second book on multicultural education, *The Light in Their Eyes: Creating Multicultural Learning Communities*, I looked more closely at the kinds of attitudes, values, and beliefs that teachers need to help develop environments where students of all backgrounds can learn effectively. In both books, I ask readers to think critically about diversity; to learn more about students in their classrooms; to develop practices that are more inclusive; and to build strong relationships with their students that help students define themselves as capable and competent learners. I hope teacher educators also find implications for what they do in both these books. But the books are not about establishing a set curriculum, or learning a specific group of strategies. In my writings, I've tried to focus on teachers' frames of mind and on teaching as an ethical endeavor. Granted, specific strategies such as cooperative education, for example, might help teachers learn their craft more effectively. But even cooperative learning, or any other pedagogical approach, can take place in the most uncooperative and oppressive of settings. And inspiring discussions can happen in the most unlikely of settings, even in classrooms with nailed-down seats in rigid rows. It isn't the strategies or approaches itself that makes the major difference, but the environments in which it takes places and the relationships on which teaching and learning are based.

NAJP: Teacher expectations are a major concern. How can teacher training best prepare teachers to be fair, neutral, impartial and objective?

SN: I would hope that all teachers would be fair and objective, but I wouldn't want them to be either "neutral" or "impartial." These words suggest a certain distance, an indifference to the students they teach, and to their backgrounds. I think this stance is problematic because it is based on the assumption that, to be fair, teachers must make believe that differences don't exist. I am reminded of one of the "golden rules" that we often invoke in teacher education, and in our society in general: that we should all be "color-blind." If being color-blind means that we treat everyone fairly and equitably, I'm all for it. But being color-blind frequently implies that differences are negative, that they need to be avoided. We are loath to hear children mention, for example, differences in skin color, or hair or food or accent, as if these things were bad.

When children mention these things we often "sshh" them up. But differences in identity are neither good nor bad, they just are, and children as well as teachers need to accept that. This doesn't mean that I'm unaware of the many and varied challenges that so many students today are facing, challenges that range from racism and exclusion, to inadequate housing, single-parent homes, and lack of English skills. These are real, and teachers need to be aware of them. But these situations cannot be used as a justification for having low expectations of students. Instead, I think that teachers need to think about the resources that students bring to their education, and to use those resources in creative and in empowering ways.

What does this mean for teacher education? I believe it means that we who are in the business of teaching future teachers need to expose them to differences of all kinds; we need to help them understand that the differences students embody can't be left outside the classroom door; and that we need to encourage them to find ways to use their students' differences in the service of their learning; it makes more sense to think about it as an asset that can be used to promote further learning. Thinking about language differences in this way seems to me a more hopeful strategy than trying to do away with them, as we have too often tried to do in our schools. The same is true of diversity in social class, ethnicity, race, and others. If we approach differences in this more hopeful way, we will have taken a giant step in encouraging teachers to have high expectations of all students.

NAJP: Using a case study approach has both positive and negative elements. Could you discuss the pros and cons of your approach?

SN: Yes, you're right, and I have always tried to keep this caution in mind. When I first toyed with the idea of doing case studies over fourteen years ago, almost everyone discouraged me because of the enormous potential that individual case studies have for perpetuating stereotypes. These warnings really helped me think about this problem more deeply, so that, when I finally wrote the case studies, I was very careful about the kinds of assertions I made about the students, their families, and their schools. I also discussed this problem upfront in the book *Affirming Diversity* itself, alerting readers repeatedly to the potential harm that case studies can cause.

At the same time, I was quick to point out how the students in the case studies also challenged the readers' stereotypes, that is, how they defied conventional expectations for their group.

For example, Avi, a Jewish student, was not always the "model student," in terms of grades, that some teachers might expect. Vinh, a Vietnamese student, claimed that he loved group work, something that most teachers might not expect of Asian students, who they believe prefer to work on their own.

On the other hand, case studies also illuminate, that is, they provide an in-depth picture of individual students and their experience, encouraging readers to think about how particular contexts of home and school can influence a student's learning. Case students can also help readers differentiate among students who might share little beyond their ethnicity. In the book, for instance, I include both a highly successful Mexican girl and a Chicano boy who had experienced a great deal of failure in school. Also, when reading case studies readers are helped to think about the kinds of challenges other students of similar backgrounds might be facing. One way I have suggested that prospective and practicing teachers gain some of the benefits of case studies is to do one themselves, and I've provided guidelines for doing so in the third edition of my book, *Affirming Diversity*.

NAJP: Teachers often do want to include multicultural education in the curriculum, but say, "there's just no time." How would you respond, and what should or can we eliminate in order to "make time"?

SN: It seems to me that teachers who respond in this way have a false notion of what multicultural education means. As I said before, curriculum integration is just one aspect of multicultural education. Multicultural education is above all a philosophy, a way of thinking about learning and teaching. To me, multicultural education is a perspective about who deserves a high quality education.

That is, multicultural education is especially about equity and access to learning. If we have a multicultural perspective, then we need to ask what I've called "profoundly multicultural questions," including: "Who is taking calculus?" "Is the bilingual program (or the special education class) in the basement?" "Who makes hiring decisions?" "Who is in the textbook?" "How are parents involved in the school?" "Do they simply have bake sales?" and so on. If they begin by asking these "profoundly multicultural questions" then my hunch is that most teachers will see that they do in fact have "time" for multicultural education. But the question itself is based on the false assumption that the knowledge that is of most worth has already been determined, and this is far from the case. Knowledge is being generated every day, and that means that what's taught is also changing every day.

We can't continue to teach our children as if the world were standing still; this includes how we define knowledge. I've often said that all knowledge is inherently multicultural; it's only when we teach it that we rob it of it's multicultural content. I remember a student once said to me, "Okay, but let's say you're teaching about whales. What's multicultural about that?" I responded, "Do you think that whales swim only off the coast of the United States?" That's just a tiny example of how we've managed to make mono-cultural what is innately multicultural.

Do teachers need to make serious decisions about what to include or take out of the curriculum? Of course they do. But believing that the major problem is that there's "not enough time" serves to downplay some of the actual underlying reasons for not including a multicultural perspective in teaching. That includes a conservative political ideology that dismisses the histories and experiences of all but Europeans and European Americans as significant enough to teach to all our children. Rather than thinking about what we need to eliminate in the curriculum in order to "make time" for multicultural education, I would suggest instead that we think about what students of all backgrounds miss when we don't make time.

NAJP: You have indicated that multicultural education is "critical pedagogy." Can you explain?

SN: If we believe that it's necessary for all students to become active and productive members of our society, then we need to help them to develop decision-making and social action skills. This also means that students need to become critical thinkers, and that they need to learn to view events and situations from a variety of perspectives. A critical pedagogy values diversity and encourages critical thinking, reflection, and action. It's opposite is what Paulo Freire calls "domesticating education" or an education that emphasizes passivity, acceptance, and submissiveness. Critical pedagogy encourages students to take risks, to be curious, to question. Rather than expecting students to repeat teachers' words, it expects them to seek their own answers through the guidance of their teachers. When I began teaching many years ago, I thought the curriculum as handed down was just "the way things were." It never occurred to me that curriculum—or any other pedagogical decisions—were influenced by people's ideologies, positions in society, experiences and power. But, I soon learned that every educational decision, made at any level, by a teacher or a school board or a superintendent, reflects the political ideology and worldview of the decision maker. This includes decisions to do away with tracking or discontinue standardized tests, lengthen the school day or use one textbook rather than another, use learning centers or rows of chairs. All of these decisions reflect particular views of learners, education, and the world. The curriculum is another example. What is excluded in the curriculum is often as telling as what is included. It's been said that history is always written by the winners, and all we need to do is look at our history books to see that this is so. This is true in all countries: one of the first things the new government did when Rhodesia became Zimbabwe, a number of years ago, was to initiate a massive project to rewrite the country's textbooks, because native Africans had been left out of their own history except in the most demeaning way by the English colonizers. In the old textbooks, Africans were little more than savages whom the English "civilized." Needless to say, the new textbooks presented history far differently. So, a question I often ask my students is "Who wrote the book?" A critical perspective helps us understand this question, and all that it implies more clearly. In our own society, the voices and perspectives of women, enslaved Africans, Native Americans, and many others have been virtually invisible in the textbooks we use.

Sometimes, their perspectives are included as a side bar or as supplementary reading, and, although the situation is changing, it is still unusual to find in our treatments of history those viewpoints that accurately reflect the position of those who haven't shared equitably in our society's benefits.

But, I'm not suggesting that multicultural perspective should simply invert the equation by substituting one "truth" or perspective for another. No, it needs to reflect on multiple and contradictory perspectives to present reality more fully. This means that we all sometimes need to learn to listen to and understand even those viewpoints with which we disagree. As a consequence, teaching with a critical multicultural perspective is not about what is "politically correct." On the contrary, having a critical perspective means opening our minds, rather than closing them. It's also very difficult. It takes a lot of time and careful thought to become an educator with a critical multicultural perspective. It means relearning a lot of what we've previously learned, and doing so with humility and generosity. It's an approach that values students' identities and respects their intelligence, so in spite of the hard work, I believe all our students—and our teachers—are worth the effort!

REFERENCES

Nieto, S. (2000). *Affirming Diversity: The Sociopolitical Context of Multicultural Education*, (3rd Ed.) New York: Longman Publishers.

Nieto, S. (1999). *The Light in their eyes: Creating Multicultural Learning Communities*. New York: Teachers College Press.

Nieto, S. (in press). *What keeps Teachers Going in Spite of Everything*. New York: Teachers College Press.

Note: The ethnic terms I use here, and that I used in my book, were chosen by the students themselves.

Straight From the Heart: Integrating Service Values With The Study Of Psychology

Judith Reifsteck
James Madison University

For the student enrolled in an undergraduate psychology class, there are important questions for debate. How do practitioners effectively treat disordered behavior in all members of the community without regard for ability to pay? Can the practitioner get at the whole problem of abnormal behavior or disordered adjustment if he or she treats only the individual in trouble, leaving social conditions unaltered and other affected members of families and communities unchanged? These issues may be addressed in the study of psychology by adding a service learning lab to the course design. As a method of addressing these goals and objectives, a four-hour service learning lab was added to two sections of a psychology of adjustment course at James Madison University. The purpose is to elucidate service values in action and to encourage the student to consider community mental health settings for future internship and employment experiences. According to surveys of the service learning lab, the experience allows students to more realistically address issues of practitioner ethics, community collaboration, and socio-cultural diversity.

A massive, wood arch stood in the kitchen hall at Summit House. The home had been a community rehabilitation center for people socially and vocationally disabled by mental illness. The visitor noticed the beautifully hand crafted keystone, doorway and arch in the middle of the room. It was brought from a different location and installed in this relatively new structure. A former client of this community service agency explained the history during a guided tour. Ten years ago, when the clubhouse moved from a small home to it's current larger facility, a committee voted to install the arch as an inspiration to future members and as a reminder of the past. The architecture committee was made up of three professional staff members and ten clients of the clubhouse. The contractor and architect remarked at the time that the plans were the most efficient and creative blueprints they had seen in many years. Client empowerment and a collaborative approach had created a unique place for these former clients because of the flexibility and compassion of the human service agency (McCluskey-Fawcett, 1992). One of the goals of

---

Author info: Correspondence should be sent to: Dr. Judith Reifsteck, Psychology Dept., James Madison University, Harrisonburg, VA. 22807. reifstjb@jmu.edu.

North American Journal of Psychology, 2002, Vol. 4, No. 3, 489-496.
© NAJP

the service learning project described in this article is to facilitate student participation in such agencies.

My own journey as a human service professional began in the 1970's in the hill country of South Texas working as a consultant to Head Start programs. Driving to pre-schools in Fredricksburg, Ingram, and Kerrville was interspersed with giving psychological tests to students attending schools in urban San Antonio and adolescents in expensive acute care or residential treatment facilities in Austin. As I traveled between public organization and private agency, one fact remained clear. Human suffering was very democratic. Regardless of the economic or social context, the families suffering and the professional's working shared the same values of compassion, collaboration, and empowerment.

Yet somehow the valuable services designed to help those in trouble do not reach all of the individuals in need, all of the time. Are we really getting at the whole problem if we treat only the individual identified for services in our human service agencies? What are the long-term consequences of leaving social conditions unaltered and other members of the effected community unchanged? How much more effective can we be if we include other members of the community and focus on socio-cultural variables that contribute to problems in adjustment? For the student enrolled in an abnormal psychology, community psychology or psychology of adjustment course, these are important questions for debate. Frequently, a curricular objective in these courses has been to define adjustment and identify the socio-cultural contexts in which it occurs (O'Sullivan, 1993; Raupp, 1992).

Successful teachers frequently struggle with the task of making factual information relevant and alive for their students (Palmer, 1998). Student assessment reveals the majority of psychology students express a preference for experiential and interactive methods of teaching in addition to lecture and a more content based approach (King, 2000; Roberts, 2000). Additionally, the undergraduate college student pursuing study in the field of psychology is faced with the objective of defining mental illness and learning more about the contemporary solutions offered as interventions for such disorders (Lewis, 1998).

Designing an approach to instruction that enhances curricular goals by adding service learning objectives is one way to partially address some of these objections regarding relevance and experience. Keeping the focus on these objectives can help students and teachers make connections to the local community and the larger social context surrounding mental health interventions (Chilcoat, 1994). There were four such objectives addressed when the current service learning experience was planned for use in a psychology class. First, students actively learned how community psychology was defined and practiced in a human service agency. Second, service learning prepared the student to begin thinking about and planning for a more extensive practicum, internship, or employment experience (Chapdelaine, 1999; Briihl, 2001). Third, service learning provided real life examples for related course work on practitioner ethics, socio-cultural context models of adjustment, and diversity in psychology (Dunlap, 1998; Miller, 1969). Fourth, the lab may be adapted to any size psychology class. Contact with community and human service agencies may be accomplished in classes large and small through field trips, various volunteer activities, and lab exercises designed by the instructor. Different teachers with different gifts and professional activities create the link to the larger community in many creative ways using the principles of socio-cultural relevance, compassion, service and collaboration as a guide.

Using student discussion as a de-briefing technique before, during and after the service project experience helped to deepen the learning and connect experience to theory. Although the time spent in assisting with agency activities was minimal, the time spent interacting in class with peers and teacher was substantial. The classroom community was used to develop student understanding of such broader concepts as community mental health research, funding challenges, de-institutionalization of the chronically mentally ill, volunteer recruitment, and training methods. Such broader discussion was unplanned and developed naturally after student exposure to the agency setting and activities.

METHOD

As a method of addressing these goals and objectives, a four hour service learning lab was added to two sections of a psychology of adjustment course in June 2000 and June 2001 at James Madison University. Students (*N* = 16) committed to three hours of preparation, reading, and class discussion and one hour of attendance at a program run by one of two selected human service agencies. For the service learning lab, the students participated in a camp for at risk children or a social clubhouse for individuals with chronic mental illness.

In class, three additional human service programs were reviewed by students in the service learning lab. They included a center for psychiatric rehabilitation, the Children's Defense Fund's summer freedom school, and the Human Service Alliance in Winston-Salem, North Carolina. Students also read and discussed two theoretical approaches which allowed them to explore themes of practitioner ethics, socio-cultural context of adjustment, and diversity in psychology. These readings described the communitarian approach to psychological discourse and action (Prilleltensky, 1997), the social movement perspective applied to the development of community crisis centers

(Echterling & Wylie, 1981), and the link between values clarification and positive personal adjustment (Kekes, 1986). These theoretical frameworks helped students make sense of their participation in human service programs. Upon completion of the lab, students filled out an evaluation survey which asked them to rate their level of satisfaction on ten items. The results of this survey indicated high levels of student satisfaction with the lab. The items were keyed to objectives of service learning in the psychology class and rated (0 points, 1 point or 2 points) on a 3 point Likert rating scale. These included 1) to help non profit agencies with the workload, 2) to improve awareness of available services, 3) to increase interest in social or political issues, 4) to challenge attitudes regarding prevention services and equitable care, 5) to increase empathy, 6) to feel more of a sense of belonging in the community, 7) to stimulate ideas for further study, internships, and career plans, 8) to increase confidence about academics or career, 9) to increase understanding of psychology, and 10) to help define career goals. The director or other key personnel from the community service agency also completed similar evaluations designed to assess progress on these objectives. These staff members gave feedback suggesting a positive impact of the project.

Students also studied the history and purpose of the community service agencies they visited. Such explorations help students to keep focused on the communitarian values of compassion, self determination, diversity, collaboration and justice which psychology has to offer through it's foundation and history as a culturally relevant institution in our society (Miller, 1969).

RESULTS

Student feedback on a program evaluation survey designed for the purposes of this lab revealed high levels of satisfaction with the project (Mean score = 17. 6 out of a possible 20 points). Satisfaction ratings from the community agencies demonstrated a positive response and a desire to continue service projects involving undergraduate college psychology students (Mean score = 18 out of a possible 20 points). The motivation for doing so included a desire for collaboration, commitment to training, and the positive impact such projects have on community programs. The students received important exposure to a community mental health setting and the service experience allowed students to more realistically address issues of practitioner ethics, community mental health, and diversity in psychology during class discussion.

Students participated in one of two service learning opportunities. One setting offered participation in a primary prevention program for at risk children. A second service learning project allowed students to become involved in a clubhouse setting for the psychosocial rehabilitation of patients with chronic mental illness. In addition, they read and discussed three selected articles on practitioner ethics and the social movement perspective in community mental health. These readings set the stage for class discussion on diversity, practitioner ethics, the socio-cultural context model of adjustment, and equitable availability of services for all mental health populations.

Students learned that effective community organizations do not simply add funds or hire designated staff as an intervention for community mental health problems. There is a developmental curve of growth and adaptation which successful community agencies follow if they seek lasting positive social change (Kilpatrick, 1996).

DISCUSSION

The communitarian approach and the social movement perspective reviewed in this service learning lab endorsed a multifaceted, environmental approach to helping people. This is an approach which is informed by underlying moral principles, grassroots advocacy, and community assessment practices. We cannot expect that expert administrators or superior programs alone will make relevant social and individual change happen (Blumer, 1969). Instead, when we are surrounded by people, organizations, and groups who articulate the values of wellness and empowerment for all individuals and families in the community, human service work holds definable value (Freund, 1993). Using this approach, the underlying moral value of the interventions and actions of psychologists can be observed and understood as socially relevant for clients in many diverse contexts.

When students and professionals in psychology collaborate on projects with an eye towards justice and fairness for all, such cooperation can improve the overall quality of our community's mental health and social well being. Teachers and practitioners who are guided by the communitarian values of caring, self-determination, diversity and distributive justice subscribe to the belief that no *one* community member will be well if *all* community members are not well. The social movement perspective explained in this service learning lab depicts agencies and organizations that learn to adapt their original purpose in order to survive change. When these changes are understood within a framework of the larger social forces of our times, the planning and delivery of services as a result of such social change can be a chaotic but exciting experience for students to observe. The students who worked in the service learning lab witnessed this in the agencies they visited. They learned that some human service programs owed their very existence and relevance to a grassroots mandate and the service values which generated the program. The advocacy and commitment of devoted individuals was part of the process of generating human benefits out of social discontent and unrest (Rachal, 1998). This was necessary in order to accomplish the social change needed to develop outreach mental health programs for troubled families who could not afford services.

Good teaching in any discipline may be conceived of as student centered, as in active learning models of education, or teacher centered, as in traditional content oriented approaches. A third model articulates values or principles at work in the course of study. When the course of study is psychology and community mental health, a roster of values may include such principles as compassion, service, self-determination, empowerment, and distributive justice. These values were defined and explicated through the research read by the students and observed in the agencies where they visited and volunteered.

One of the students in the service learning lab spoke of the work he did with a group of children in the following way: "...as we played a game, I felt a moment of comfort knowing that today I was working to meet the needs of teens and children who would otherwise be at home alone engaged in doing nothing or worse. Over fifty children received for the first time that summer what other children take for granted if they have it. The chance to form friendships, to learn, to read, to play, and to relax under the watchful eye of caring adults. Some of the people I saw working at this camp were volunteering, others worked for pay, but it was not money that brought us here. It was caring."

Like the arch over the door in the hall at Summit House, community programs like this one can inspire and remind students of the relevance of the principles they learn in psychology class. Caring is the keystone of the arch. It holds the pillar of social change on one side, buttressed against the pillar of individual growth on the other.

REFERENCES

Blumer, H. (1969). Elementary collective groupings. In A. McClung Lee (Ed.), *Principles of Sociology*. New York. Barnes & Noble.

Briihl, D. (2001). Life after college: Psychology students' perceptions of salary, business hiring criteria, and graduate admission criteria. *North American Journal of Psychology*, *3*, 321-330.

Chapdelaine, A., & Chapman, B.L. (1999). Using community based research projects to teach research methods. *Teaching of Psychology, 26*, 101-105.

Chilcoat, G. & Ligon, J. (1994). Developing democratic citizens: The Mississippi freedom schools as a model for social studies instruction. *Theory and Research in Social Education, 22*, 128-175.

Dunlap, M.R. (1998). Methods of supporting students' critical reflection in courses incorporating service learning. *Teaching of Psychology, 25*, 208-210.

Echterling, L. & Wylie, M.L. (1981). Crisis centers: A social movement perspective. *Journal of Community Psychology, 9*, 342-346.

Freund, P. (1993). Professional roles in the empowerment process: Working with people with a psychiatric disability. *Psychosocial Rehabilitation Journal, 16*, (3), 65-73.

Kekes, J. (1986). The informed will and the meaning of life. *Philosophy and Phenomenological Research, 67*, (1), 65-72.

Kilpatrick, J. & Danzinger, S. (1996). *Better than money can buy: The new volunteers*. Winston-Salem, N.C.. InnerSearch Publishing.

King, R. (2000). Portfolio development: Using authentic learning assignments in psychology courses. *North American Journal of Psychology, 2*, 151-166.

Lewis, J., Lewis, M., Daniels, J. & D'Andrea, M. (1998). *Community counseling: Empowerment strategies for a diverse society*. Pacific Grove, CA.: Brooks Cole.

McCluskey-Fawcett, K., & Green, P. (1992). Using community service to teach developmental psychology. *Teaching of Psychology, 19*, 150-152.

Miller, G. A. (1969). Psychology as a means of promoting human welfare. *American Psychologist, 24*, 1063-1075.

O'Sullivan, M.J. (1993). Teaching undergraduate community psychology: Integrating the classroom and the surrounding community. *Teaching of Psychology, 20*, 80-83.

Palmer, P. (1998). *The courage to teach: Exploring the inner landscape of a teacher's life*. San Francisco, CA: Jossey Bass.

Prilleltensky, I. (1997). Values, assumptions and practices: Assessing the moral implications of psychological discourse and action. *American Psychologist, 52*, 517-535.

Rachal, J. (1998). We'll never turn back: Adult education and the struggle for citizenship in Mississippi's Freedom Summer. *American Educational Research Journal, 35*, 167-198.

Raupp, C. D., & Cohen, D.C. (1992). "A thousand points of light" illuminate the psychology curriculum: Volunteering as a learning experience. *Teaching of Psychology, 19*, 25-30.

Roberts, C. (2000). The efficacy of speech instruction and experience in an upper division psychology course. *North American Journal of Psychology*, 2, 139-143.

APPENDIX A

STUDENT SURVEY OF SERVICE LEARNING LAB

Please circle the answer that best applies to your experience with the service learning lab in this course. (0 = not at all true, 1 = somewhat true, 2 = very much true)

The service learning lab........

1. Helped me define my career goals	0	1	2
2. Increased my understanding of psychology	0	1	2
3. Was more challenging than other projects I've had	0	1	2
4. Allowed me to integrate material from class to a real life situation.	0	1	2
5. Increased my interest in social or political issues	0	1	2
6. Changed my attitude on some issues	0	1	2
7. Increased my confidence about academics or career choice	0	1	2
8. Increased my empathy as a person	0	1	2
9. Helped me feel more a sense of belonging in my community	0	1	2
10. Stimulated ideas about further study, internship plans or career plans	0	1	2

APPENDIX B

STAFF SURVEY OF SERVICE LEARNING LAB

Please circle the answer that best applies to your experience with the students and programming from the service learning lab in this course (0 = not at all true, 1 = somewhat true, 2 = very much true)

The service learning lab........

1. Helped our agency with the work load	0	1	2
2. Improved community awareness of our services	0	1	2
3. Was the right amount of work for the benefits we received	0	1	2
4. Allowed our agency to fulfill the mission of training future staff	0	1	2
5. Led to a beneficial collaboration with new friends	0	1	2
6. Increased positive attitudes on some issues	0	1	2
7. Increased confidence about the agency's mission	0	1	2
8. Increased my empathy as a person	0	1	2
9. Helped me feel more of a sense of belonging in my community	0	1	2
10. Stimulated ideas about further programs, internship plans or program plans	0	1	2

How to Submit Papers to *NAJP*

The submission of good papers is very important to any psychology journal. We hope that you will support our views about what a psychology journal should be like by submitting papers now. We will try to mail a decision on your paper within 5 weeks of receipt by us. We will publish your paper in an average of 5 months of receipt of a final accepted version.

Submit 1 paper copy of each manuscript prepared according to APA guidelines. Also include a diskette with your manuscript or (preferably) e-mail the file to NAJP@mpinet.net. Include 2 large stamped envelopes so we can send your manuscript out for review. We encourage you to have your paper read by peers before submission.

If the final paper is accepted, please send a check or purchase order for the amount specified with your final draft (billed at $25 dollars per typeset page in units of two pages).

Submission of a manuscript to the Journal is a contract to publish with the Journal if the paper is accepted. Because of our rapid editorial feedback, simultaneous submission to other journals is not allowed. We will offer brief comments on the suitability of a topic for publication in *NA*JP before a submission.

Send all correspondence to: *North American* Journal of Psychology, 240 Harbor Dr., Winter Garden, FL 34787, phone 407-877-8364.

SUBSCRIPTIONS

All subscriptions are on a calendar year basis only. If you order late in the year, the earlier issues will be sent to you. Institutional rate, $65. Individual rates, paid by check, $30. Add $7 for Canadian subscriptions. Add $8 for subscriptions outside North America. Phone orders accepted with Visa, MasterCard or American Express. Call 407-877-8364. The *North American* Journal of Psychology is now published semi-annually in June & December.

Visa ☐ AMEX ☐ MC ☐

Card # Exp. Date

Name

Address

City State Zip

A Summary of Our Philosophy

The ***North American*** **Journal of Psychology** publishes scientific papers that are of general interest to psychologists and other social scientists. Papers may encompass a wide variety of topics, methods, and strategies. It is our belief that many good scientific papers are rejected for invalid reasons. Some journals reject up to 90% of their submissions because they don't have space. Furthermore, studies of the review process strongly suggest that not all of the lucky 10% are better papers than those that get rejected. In our view many of those rejected papers are meritorious with need for only minor revisions. Perhaps our philosophy can be summed up this way: we are looking for ways to strengthen good papers rather than searching for weak excuses to reject them.

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

ERIC®

NOTICE

Reproduction Basis

☑ This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

☐ This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").

EFF-089 (3/2000)